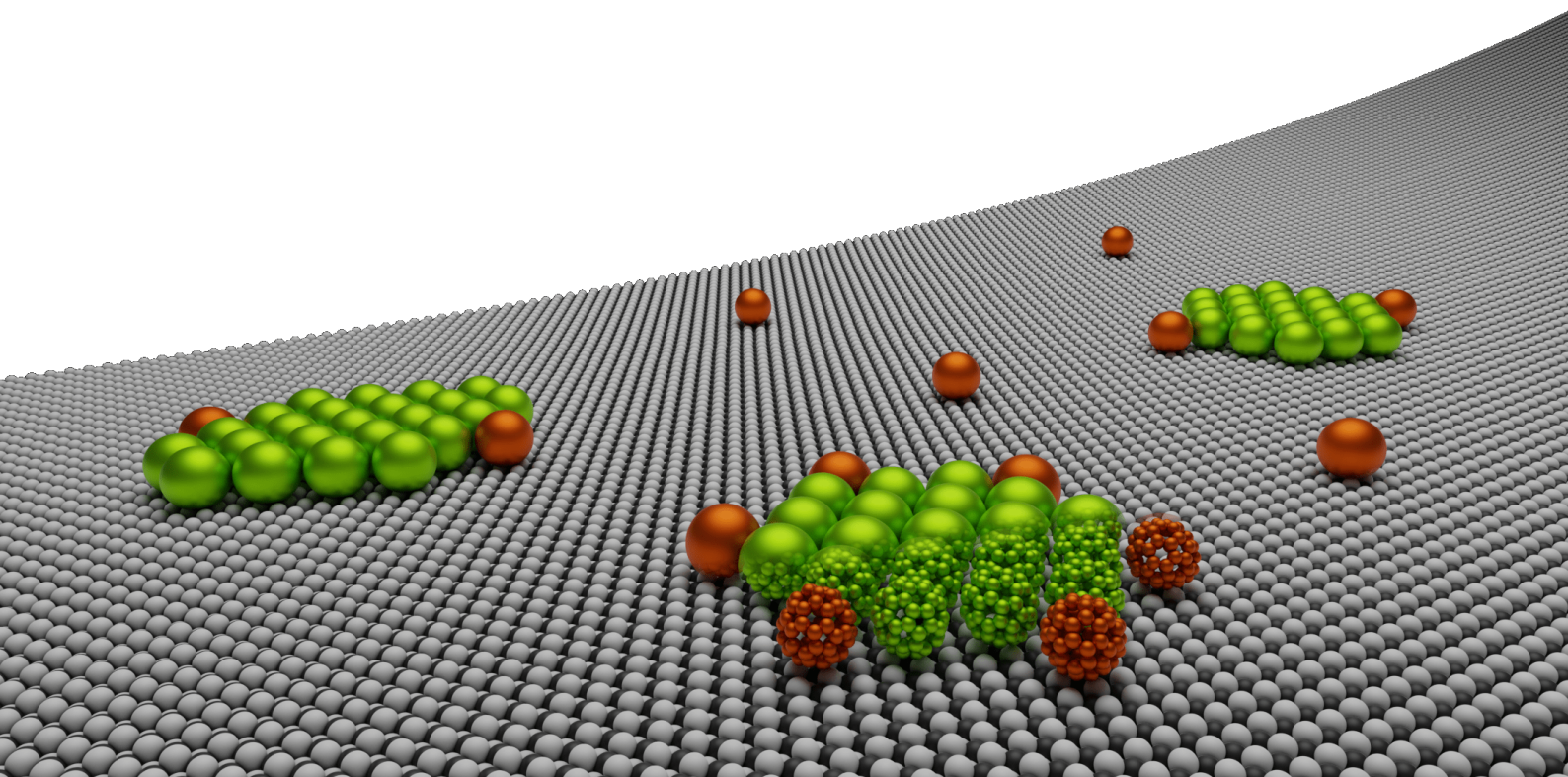


# MULTI-SCALE MODELLING OF THE EPITAXIAL GROWTH OF ORGANIC THIN FILMS ON INSULATING SURFACES

DISSERTATION  
ZUR ERLANGUNG DES GRADES  
"DOKTOR DER NATURWISSENSCHAFTEN"  
AM FACHBEREICH PHYSIK, MATHEMATIK UND INFORMATIK  
DER JOHANNES GUTENBERG-UNIVERSITÄT  
IN MAINZ

WILLIAM JANKE  
GEB. IN MEISENHEIM AM 04.11.1989  
MAINZ, DEN 11.08.2021



*Multi-Scale Modelling of the Epitaxial Growth of Organic Thin Films on Insulating Surfaces*

LOCATION:  
Mainz

DATE OF DEFENSE:  
December 7th 2021

## LIST OF PUBLICATIONS

This thesis is in part a recapitulation of material published by the author during the PhD program in the following articles:

- [1] **W. Janke** and T. Speck, "Modeling of epitaxial film growth of  $C_{60}$  revisited". In: *Phys. Rev. B* 101.12, 125427 (2020)

*Own contribution:* I set up the molecular dynamics simulations (using LAMMPS), calculated transition rate parameters from the data, and made the modelling of the transition rates. Moreover, I wrote the kinetic Monte Carlo simulation code, performed the simulations and analyzed the data. I created all the figures and wrote most of the manuscript.

- [2] **W. Janke**, T. Speck, F. Loske, J. Lübke, J. Schütte, M. Reichling, and A. Kühnle, "Erratum: Quantitative description of  $C_{60}$  diffusion on an insulating surface [*Phys. Rev. B* 82, 155428 (2010)]". In: *Phys. Rev. B* 101.4, 049907 (2020)

*Own contribution:* I found the error during testing of my kinetic Monte Carlo simulation code and wrote the erratum for A. Kühnle upon request.

- [3] **W. Janke** and T. Speck, "Multiscale modeling of structure formation of  $C_{60}$  on insulating  $CaF_2$  substrates". In: *J. Chem. Phys* 154.23, 234701 (2021)

*Own contribution:* I derived the analytic expression for the used interaction potentials, set up the molecular dynamics simulations (using LAMMPS), analyzed the data and made the transition rate modelling. I created all the figures and wrote most of the manuscript. I also designed the rendering that was featured on the cover of this issue.



# ABSTRACT

Growth morphologies of epitaxially grown molecular thin films are of interest for a range of different technological applications. While the possibilities of epitaxial growth experiments are usually explored in the lab, it is desirable to be able to accurately simulate the process to more quickly explore a wide range of different experimental protocols. The kinetic Monte Carlo (KMC) method is a promising approach in that regard, as it can access the necessary time and length scales on which epitaxial growth can be observed. However, a KMC algorithm requires the transition rates of all implemented elementary transitions as input parameters and ideally also their dependence on experimental parameters like temperature. Experimental data can help to inform a model for the transition rates of a KMC simulation, but experiments alone are not sufficient to completely determine all the parameters. As a result, KMC rate models are often oversimplified to be able to work with a limited amount of experimental information. In this thesis, we present a bottom-up approach for the determination of a KMC rate model that is built on a foundation of transition rate data gathered in molecular dynamics (MD) simulations. The example system on which we apply this approach is the epitaxial growth of the buckminsterfullerene  $C_{60}$  on a calcium fluoride substrate ( $CaF_2(111)$ ). We set up MD simulations of this system in a wide variety of configurations in which we can observe the elementary transitions, determine their transition rates in a range of temperatures and finally use the obtained data to derive a rate model for use in KMC simulations. To test the obtained models, we run KMC simulations and compare the results with experimental data. This thesis contributes to the ability to model and simulate the self-assembly processes of molecules on insulating substrates. Consequently, it advances the understanding of such systems and enables the development of new strategies to control the evolution of cluster morphologies in deposition experiments.



# ZUSAMMENFASSUNG

Wachstumsmorphologien epitaktisch gewachsener molekularer Dünnschichten sind für eine Reihe unterschiedlicher technologischer Anwendungen von Interesse. Während die Möglichkeiten epitaktischer Wachstumsexperimente normalerweise im Labor untersucht werden, ist es wünschenswert, den Prozess genau simulieren zu können, um eine breite Palette verschiedener experimenteller Protokolle schneller untersuchen zu können. Die kinetische Monte Carlo (KMC) Methode ist in dieser Hinsicht ein vielversprechender Ansatz, da sie auf die notwendigen Zeit- und Längenskalen zugreifen kann, auf denen epitaktisches Wachstum stattfindet. Ein KMC Algorithmus benötigt jedoch als Eingangsparameter die Übergangsraten aller implementierten Elementarübergänge und idealerweise auch deren Abhängigkeit von experimentellen Parametern wie der Temperatur. Experimentelle Daten können in die Modellierung der Übergangsraten einer KMC-Simulation mit einfließen, aber Experimente allein reichen nicht aus, um alle Parameter einer KMC Simulation vollständig zu bestimmen. Infolgedessen werden KMC Ratenmodelle oft stark vereinfacht, um mit der begrenzten Menge an experimentellen Informationen arbeiten zu können. In dieser Dissertation präsentieren wir einen "Bottom-up" Ansatz zur Bestimmung eines KMC Ratenmodells, das auf einer Grundlage von berechneten Übergangsraten aus Molekulardynamiksimulationen (MD) aufbaut. Das Beispielsystem, auf das wir diesen Ansatz anwenden, ist das epitaktische Wachstum des Buckminsterfullerens  $C_{60}$  auf einem Kalziumfluorid Substrat ( $CaF_2(111)$ ). Wir setzen MD Simulationen dieses Systems in einer Vielzahl von Konfigurationen an, in denen wir die elementaren Übergänge beobachten, ihre Übergangsraten in einem Temperaturbereich bestimmen und schließlich aus den erhaltenen Daten ein Ratenmodell für den Einsatz in KMC-Simulationen ableiten. Um die erhaltenen Modelle zu testen, führen wir KMC Simulationen durch und vergleichen die Ergebnisse mit experimentellen Daten. Diese Arbeit trägt dazu bei, die Selbstorganisationsprozesse von Molekülen auf isolierenden Substraten zu modellieren und zu simulieren. Folglich fördert sie unser Verständnis solcher Systeme und ermöglicht die Entwicklung neuer Strategien zur Kontrolle der Evolution von Clustermorphologien in epitaktischen Wachstumsexperimenten.





# CONTENTS

I	INTRODUCTION	1
II	THEORY & METHODS	5
II.1	Experimental Details	6
II.1.1	Molecular Beam Epitaxy (MBE)	6
II.1.2	Atomic Force Microscopy	7
II.1.3	System of Interest: C <sub>60</sub> on CaF <sub>2</sub> (111)	9
II.2	Epitaxial Cluster Growth	12
II.2.1	Elementary Diffusive Transitions	12
II.2.2	Cluster Shapes	16
II.2.3	Cluster Densities	18
II.3	Kinetic Monte Carlo	21
II.3.1	Infrequent-Event Systems	21
II.3.2	Rejection-Free KMC	23
II.3.3	Transition Rate Modelling	25
II.3.4	Harmonic Transition State Theory	28
II.3.5	Transition rate measurements in MD Simulations	31
II.4	Molecular Dynamics	34
II.4.1	Equations of Motion	34
II.4.2	Velocity Verlet Integration	36
II.4.3	Periodic Boundary Conditions	37
II.4.4	Initialization and Thermostating	38
II.4.5	Generic Simulation Procedure	41
III	C <sub>60</sub> ON C <sub>60</sub> (111)	43
III.1	MD Simulations	45
III.1.1	The Girifalco Potential	45
III.1.2	Simulation Setup	48
III.1.3	Raw Transition Rate Measurements	49
III.1.4	Transition Rate Modelling and Discussion	52
III.2	KMC Simulations	56
III.2.1	Tuning of the "Simple" model	56
III.2.2	Multilayer Growth	61
III.2.2.1	Grain Boundaries	64

III.3	Summary . . . . .	66
IV	$C_{60}$ ON $CaF_2(111)$	69
IV.1	MD Simulations . . . . .	71
IV.1.1	Modelling and Interaction Potentials . . . . .	71
IV.1.2	Free Diffusion . . . . .	75
IV.1.2.1	Minimum Energy Paths . . . . .	76
IV.1.2.2	Energy Landscapes . . . . .	78
IV.1.2.3	Diffusion Coefficients . . . . .	80
IV.1.2.4	Summary and Discussion . . . . .	83
IV.1.3	Transition Rate Modelling . . . . .	86
IV.1.4	Edge Diffusion . . . . .	89
IV.1.4.1	One Initial Neighbour . . . . .	90
IV.1.4.2	Two Initial Neighbours . . . . .	91
IV.1.5	Summary . . . . .	96
IV.2	KMC Simulations . . . . .	98
IV.2.1	Cluster Densities . . . . .	99
IV.2.2	Cluster Morphologies . . . . .	104
V	PERSPECTIVES	109
A	APPENDIX: KMC IMPLEMENTATION	113
A.1	Implemented Geometry and Framework . . . . .	114
A.1.1	Lattice Coarse-Graining . . . . .	114
A.1.2	The GridNode Class . . . . .	116
A.1.3	The RateModel Class . . . . .	117
A.1.4	The Input Class . . . . .	118
A.1.5	The FullereneKMC Class . . . . .	119
A.2	Fast Diffusion . . . . .	120
A.3	Performance . . . . .	126
A.3.1	System size scaling . . . . .	126
A.3.2	Temperature scaling . . . . .	128
B	APPENDIX	131
B.1	Pseudorandom Number Generators . . . . .	132
B.2	Quaternion approach to obtain optimal rotations . . . . .	133
B.3	Properties of the Exponential Distribution . . . . .	135
B.4	Example LAMMPS input script . . . . .	140
B.5	Acknowledgements . . . . .	161

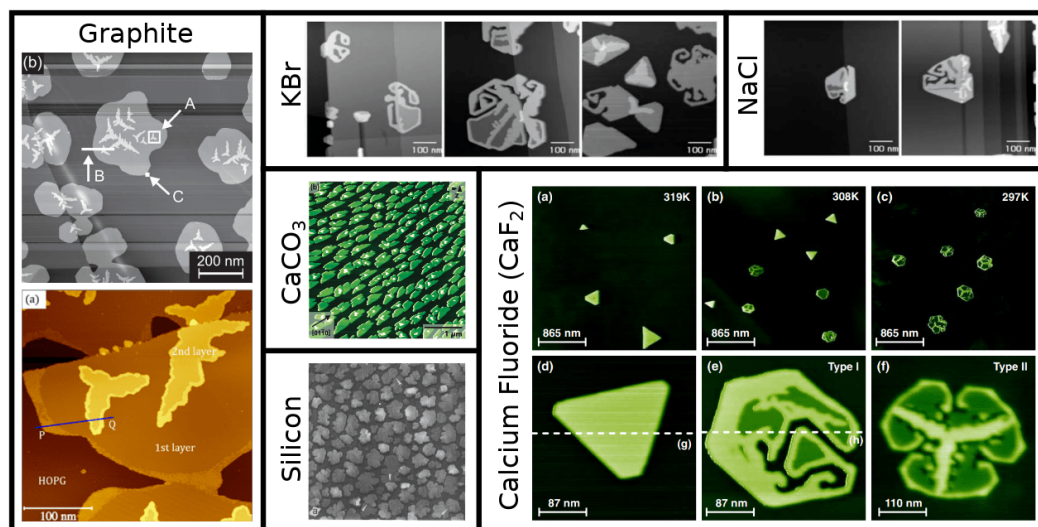
# I

## INTRODUCTION

The self-assembly of organic molecules on insulating or metallic substrates is used and researched for a variety of technological applications. Prominent examples are the development of organic photovoltaics [4] and the field of molecular electronics, where self-assembled monolayers can be used to manufacture electronic devices at a molecular level [5]. Seen as a possible gateway to such devices, the research on the growth processes of organic molecules on insulating [6] or metallic [7, 8] substrates has sustained a decent amount of interest. The large number of possible adsorbate-substrate combinations enables the ability to tune the molecule-surface interaction, resulting in the evolution of a range of different interesting molecular structures [9–12].

One adsorbate molecule that is the focus of many studies is the fullerene  $C_{60}$ . The extensive amount of research done after its discovery and first synthesis in the mid-80s [13] - ranging from the observation of its crystal growth [14, 15] over the discovery of structural phase transitions in bulk [16–18] and in thin films [19–21] to the development and study of coarse-grained  $C_{60}$ - $C_{60}$  interaction potentials like the Girifalco potential [22, 23] - have resulted in an extraordinary understanding of its physical properties. In turn,  $C_{60}$  has become a paradigmatic choice as an adsorbate for deposition experiments on metallic [24–29] and insulating [30–48] substrates with an impressive variety of resulting cluster morphologies (See Fig. 1), affected by the choice of substrate, substrate temperature and incoming adparticle flux.

However, even for an intensely studied particle like  $C_{60}$ , well-founded structure prediction and design principles for self-assembly are largely missing. A pathway to improvement is the application of computer simulation methods. As the large length and time scales necessary for the observation of epitaxial growth (system sizes spanning multiple micrometres and deposition/relaxation phases that can take several minutes or hours) are way out of range for conventional



**Figure 1:** Examples of cluster growth experiments involving deposition of  $C_{60}$  on different insulating substrates, showing a variety of different growth morphologies. The images are reprinted from various publications: The top left graphite image is reprinted with the permission of IOP Publishing from Ref. [42]; permission conveyed through Copyright Clearance Center, Inc.. The bottom left graphite image is reprinted with permission from Ref. [45]. Copyright 2018 American Chemical Society. The KBr and NaCl images are reprinted with permission from Ref. [35]. Copyright 2007 by the American Physical Society. The  $CaCO_3$  image is reprinted with permission of the Royal Society of Chemistry from Ref. [38]; permission conveyed through Copyright Clearance Center, Inc.. The silicon image is reprinted from Ref. [30], Copyright 1994, with permission from Elsevier. The  $CaF_2$  images are reprinted with permission from Ref. [53]. Copyright 2011 by the American Physical Society.

molecular dynamics (MD) methods, one typically resorts to the kinetic Monte Carlo (KMC) method [49] (also known as Gillespie algorithm[50, 51]), which has proven to be able to simulate the cluster growth of several deposition experiments [52–58].

While KMC simulations hold enormous potential, they do, however, also come with a major challenge as they require a model for all the possible elementary transition rates of the simulated system. This commonly results in models with a lot of free parameters or sometimes even thermodynamic inconsistencies. Ideally, one would like to base a KMC rate model on direct measurements of transition rates in experiments. However, measuring transition rates of the elementary transitions of epitaxial growth can be extremely difficult [59], even for the most basic transition types (like free diffusion or edge diffusion). Therefore, the construction of a complete rate model based on experimental measurements seems out of reach at the moment and the application of computational methods is necessary.

In this thesis, we are going to tackle the problem by establishing a bottom-up approach for the determination of all relevant rates for a KMC model with a minimal amount of free parameters and with considerations of thermodynamic consistency based on MD simulations of elementary transitions. We are going to apply this approach on the example of  $C_{60}$  epitaxial growth on  $CaF_2(111)$ , which has shown rich structural behaviour [36, 53, 60], captured using non-contact atomic force microscopy (NC-AFM) imaging. At room temperature, it undergoes molecular dewetting (transfer of molecules from the base layer of a cluster to higher layers [61]) and forms interesting two-layered cluster morphologies. The available data from these experiments provide a solid data set on which a KMC simulation can be tested.

The dissertation is structured in the following way: We start in chapter II with the theoretical background and employed methods. We provide some details on epitaxial growth experiments, discuss some theoretical concepts of epitaxial growth and introduce the two simulation methods that we employ in this thesis (KMC and MD simulations).

In chapter III, we show results of MD simulations of  $C_{60}$  molecules (interacting via the Girifalco potential) in which we measure elementary transition rates to derive a rate model for the epitaxial growth of  $C_{60}$  on  $C_{60}(111)$  (applicable to the second and higher layers in multilayer growth simulations). We show that the raw transition rates that we measure in the MD simulation lead to thermodynamic inconsistencies in a KMC simulation if they are adopted without adjustments and we provide an adjusted "Simple" model that is thermodynamically consistent. The obtained model is then tested against data from multilayer growth experiments of  $C_{60}$  on mica. The content of this chapter is in large part a recapitulation of our first publication (Ref. [1]).

In chapter IV, we apply the same approach to  $C_{60}$  on  $CaF_2(111)$ . As there are no established interaction potentials for this system, we derive (coarse-grained and atomistic) interaction potentials for  $C_{60}$  with  $CaF_2$  with one free parameter ( $\epsilon_F$ ), which determines the interaction strength between the adsorbate and the substrate. We present MD simulations employing both the atomistic and coarse-grained models, analyze and compare the results and finally derive rate models based on the measured transition rates. We then combine the rate models of chapters III and IV to run KMC simulations of the multilayer cluster growth of  $C_{60}$  on  $CaF_2(111)$ . We compare the results to experimentally observed cluster densities and morphologies. While the first part of this chapter (MD simulations and modelling) is mostly a recapitulation of our second publication (Ref. [3]), the presented KMC simulation results are unpublished as of the submission date of this dissertation.

Details on our implementation of the KMC algorithm are given in the appendix chapter [A](#). Miscellaneous additional information is provided in the appendix chapter [B](#).

# II

## THEORY & METHODS

### CONTENTS

---

iI.1	Experimental Details . . . . .	6
iI.1.1	Molecular Beam Epitaxy (MBE) . . . . .	6
iI.1.2	Atomic Force Microscopy . . . . .	7
iI.1.3	System of Interest: $C_{60}$ on $CaF_2(111)$ . . . . .	9
iI.2	Epitaxial Cluster Growth . . . . .	12
iI.2.1	Elementary Diffusive Transitions . . . . .	12
iI.2.2	Cluster Shapes . . . . .	16
iI.2.3	Cluster Densities . . . . .	18
iI.3	Kinetic Monte Carlo . . . . .	21
iI.3.1	Infrequent-Event Systems . . . . .	21
iI.3.2	Rejection-Free KMC . . . . .	23
iI.3.3	Transition Rate Modelling . . . . .	25
iI.3.4	Harmonic Transition State Theory . . . . .	28
iI.3.5	Transition rate measurements in MD Simulations . . . . .	31
iI.4	Molecular Dynamics . . . . .	34
iI.4.1	Equations of Motion . . . . .	34
iI.4.2	Velocity Verlet Integration . . . . .	36
iI.4.3	Periodic Boundary Conditions . . . . .	37
iI.4.4	Initialization and Thermostating . . . . .	38
iI.4.5	Generic Simulation Procedure . . . . .	41

---

## II.1 EXPERIMENTAL DETAILS

While we are not conducting any experiments ourselves in this thesis, we are interested in reproducing past and future epitaxy experiments with our simulation techniques. Without going into much detail, we are going to discuss a few experimental details in this section to provide a basic understanding of the experiments of interest.

### II.1.1 MOLECULAR BEAM EPITAXY (MBE)

Molecular beam epitaxy (MBE) is one of the fundamental tools of nanoscience [62] and has been used to study the epitaxial growth of thin films on surfaces since the 60s [63, 64]. "Molecular beam" describes a stream of particles that is ejected from a device through an ultra-high vacuum (UHV) with negligible interaction with itself or other gas particles. The term "epitaxy" (first introduced in 1928 [65]) stands for the crystalline layer growth upon ("epi") deposition onto a crystalline substrate, the crystalline orientation of which imposes an order ("taxis") on the orientation of the grown layers. One can distinguish between "homoepitaxy" if the deposit and the substrate are made of the same material and "heteroepitaxy" if the materials are different [66]. While the surrounding apparatus of MBE can be very involved, the basic setup is easily sketched as shown in Fig. 2.

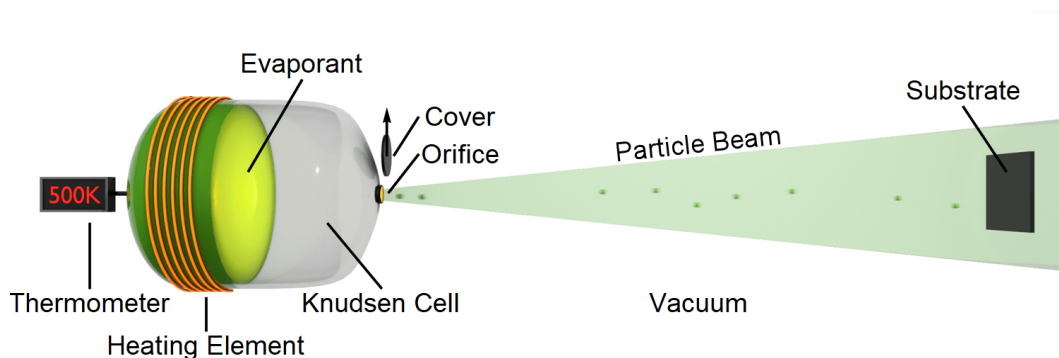
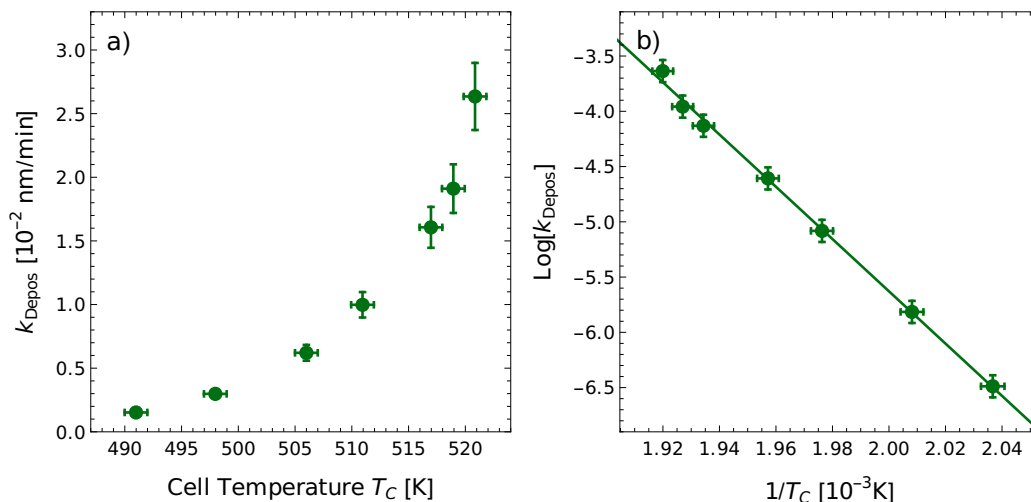


Figure 2: Sketch of a molecular beam epitaxy setup.

A so-called "Knudsen Cell" houses a probe of the evaporant material that is to be deposited on a substrate. A heating element heats the probe to the point of sublimation and the resulting particle gas can be released through an orifice by removing a cover. A UHV has to be sustained for the particle beam to not get disturbed and for the substrate to stay as free of impurities as possible. The particle flux can be controlled by changing the temperature of the evaporant via the heating element. To monitor the temperature of the





**Figure 3:** Deposition rate of C<sub>60</sub> molecules from a Knudsen cell on a substrate plotted against cell temperature. a) Normal plot. b) Arrhenius plot, showing that the data follows an Arrhenius type behaviour (Eq. II.1). The values are taken from Ref. [60].

evaporant a thermometer can be placed in the vicinity of the material. The evaporation rate – and consequently the deposition rate – follows an Arrhenius behaviour

$$k_{\text{Depos}} \sim e^{\frac{-E_S}{k_B T}} \quad (\text{II.1})$$

with Boltzmann constant  $k_B$ , temperature  $T$  and sublimation enthalpy  $E_S$ . In Fig. 3 an example of this behaviour is shown with the deposition rate of C<sub>60</sub> molecules from a Knudsen cell against the cell temperature. This information can in principle be used to set a wide range of desired deposition rates in the experiment. However, in practice this is hard for experimentalists to do as the deposition rate of any given Knudsen cell decays over time and Knudsen cells are often made in house without precise manufacturing techniques.

Upon deposition on the substrate, the molecules (also called adparticles) diffuse around on the surface, nucleating into clusters that can then grow and coalesce into thin layers.

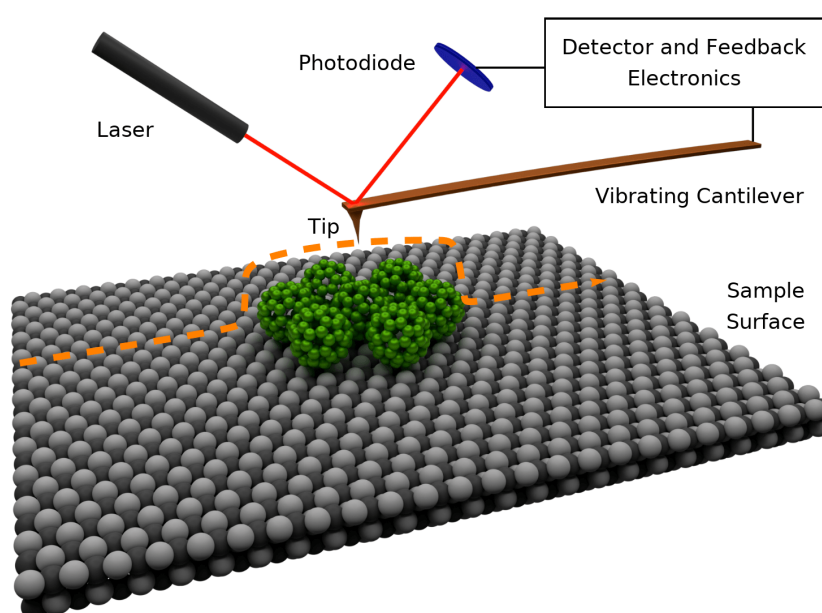
## II.1.2 ATOMIC FORCE MICROSCOPY

While the invention of the scanning tunnelling microscope in 1982 [67] quickly made atomic resolution imaging of conducting surfaces possible [68] via the measurement of a tunnelling current between a conducting tip and a conducting surface, non-conducting surfaces like insulators evaded successful imaging for a longer period. During the development of the STM, forces between the tip and

the sample have been observed, spawning the idea of atomic force microscopy (AFM) in 1986 [69]: measuring those tip-sample forces to image the sample surface. However, making an AFM work for atomic resolution imaging turned out to be a harder problem and it took the development of frequency modulated non-contact atomic force microscopy (NC-AFM) [70] and UHV conditions to achieve first atomic resolution images with an AFM almost ten years after its invention [71, 72], opening up the possibility of imaging any kind of non-conducting surface. Improvements in the technology later enabled imaging of subatomic features [73, 74], chemical bonds [75] and even the probing of the interaction force between a single pair of  $C_{60}$  molecules [76].

While details of frequency-modulated NC-AFM fall way out of the scope of this short introductory section (we refer to the extensive book series by Morita et al. [77–79] for details), we are providing a basic picture of this technology. A sketch of a NC-AFM setup is shown in Fig. 4. A sample surface is brought into proximity of a cantilever tip via piezoelectric elements. The cantilever is excited into a vibration that can be observed by reflecting a laser off of the cantilever’s back onto a photosensitive diode (PD). The feedback electronics use the PD signal to excite the cantilever on its resonance frequency  $f_0$ . While the cantilever itself can be understood as a damped harmonic oscillator, the presence of a sample surface in close proximity of the cantilever tip adds a perturbation to the oscillator via the acting van der Waals forces between the tip and the sample. This perturbation results in a shift of the resonance frequency to  $f = f_0 + \Delta f$ , which changes depending on the tip-sample distance. The detector electronics can pick up this change in the resonance frequency,

Figure 4: Sketch of an NC-AFM setup.



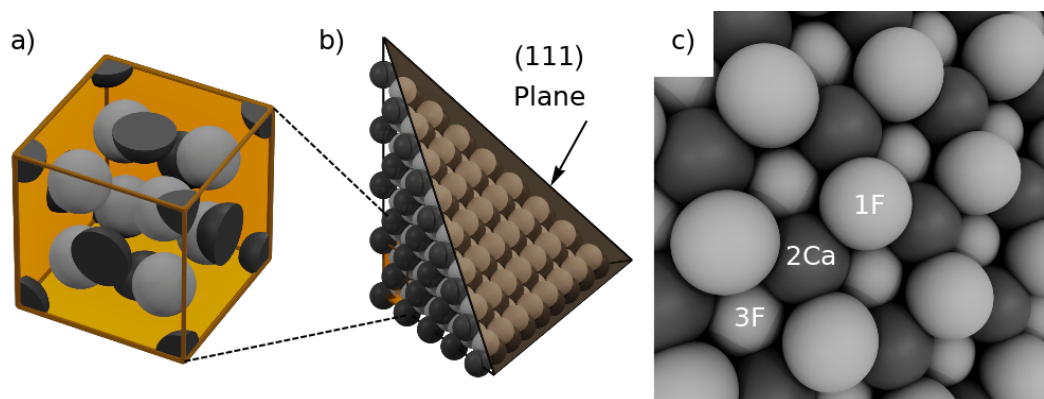
thereby indirectly measuring the tip-sample distance / the tip-sample force. To map out the sample surface, it is moved beneath the cantilever tip, either in "constant interaction mode" where the distance between the tip and the sample is kept constant via a feedback loop or in "constant height mode" where the probe is not moved along the z-axis during measurement and the change of the interaction strength is recorded.

### II.1.3 SYSTEM OF INTEREST: $C_{60}$ ON $CaF_2(111)$

As already mentioned in the introduction, there surely are many systems to choose from when it comes to deposition of  $C_{60}$  molecules on different kinds of surfaces, most with interesting properties of their own. However, in this work, we are mainly focusing on the epitaxial growth of  $C_{60}$  on an insulating  $CaF_2(111)$  surface.

The structure of calcium fluoride and its (111) surface plane are shown in Fig. 5(a-c). The arrangement of calcium and fluoride atoms in the unit cell [Fig. 5(a)] is also known as "fluorite structure". A close-up of the (111) surface layer is shown in Fig. 5(c), exposing a triple-layered structure, where a layer of calcium (2Ca) is placed in between two layers (1F and 3F) of fluoride, each layer being organized in a hexagonal lattice. At cleavage, a separation between adjacent triple-layers (i.e. separation between two layers of fluoride) is energetically favourable, which always leads to  $CaF_2(111)$  surfaces being terminated by a layer of fluoride. If done under UHV conditions to prevent degradation through air exposure [80], the right cleavage techniques [81] can prepare large atomically flat terraces extending over several  $\mu m^2$ .

An atomistic visualization of  $C_{60}$  molecules on the  $CaF_2(111)$  surface is shown in Fig. 6. The  $C_{60}$  molecule consists of 60 C atoms arranged in a soccer ball



**Figure 5:**  $CaF_2(111)$  structure. Calcium atoms are shown in grey, fluoride atoms in light grey. (a) Unit cell of  $CaF_2$ . (b)  $CaF_2$  crystal cleaved along the (111) plane. (c) Close-up of the  $CaF_2(111)$  surface, revealing its triple-layer structure.

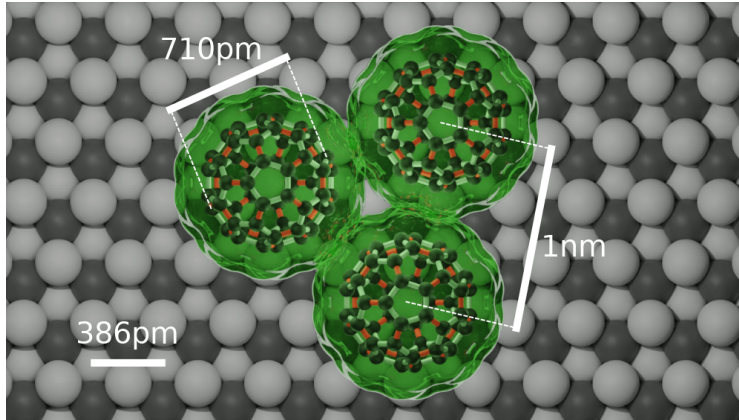
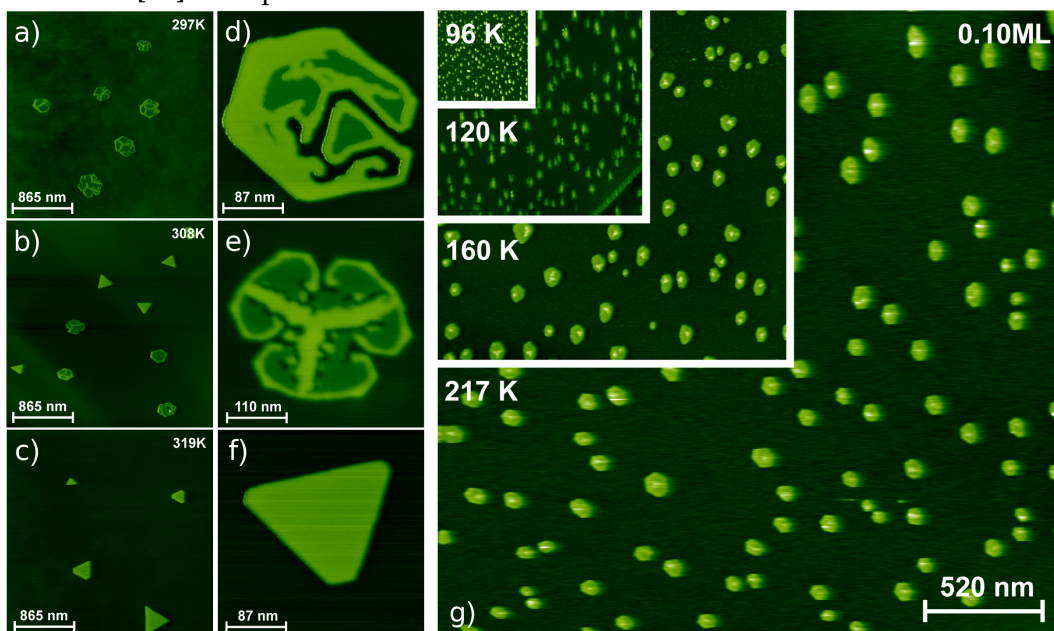


Figure 6: Atomistic visualization of  $C_{60}$  on  $CaF_2(111)$ .

like spherical structure. The nucleus-to-nucleus diameter of  $C_{60}$  is roughly  $2R = 0.71$  nm while the Van der Waals diameter is close to  $2R_{vdW} = 1$  nm. The illustration in Fig. 6 shows nicely, how the  $C_{60}$  molecules can cluster with negligible lattice strain on the  $CaF_2(111)$  surface despite the different lattice parameters. In experiments, this results in very regularly shaped hexagonal or triangular clusters upon deposition.

Figure 7: NC-AFM images from experiments of  $C_{60}$  clusters after deposition on  $CaF_2(111)$  at several different temperatures. (a,b,c) large scale NC-AFM images of multiple clusters, showing the triangular-shaped two-layered cluster at high temperatures and the mostly one-layered hexagonal shapes with complex branched structures and outer rims in the second layer at room temperature. (d,e,f) Close-ups of a few cluster morphologies at and above room temperature. (g) Cluster density measurements at low temperatures. Reprinted with some adjustments from Ref. [60] with permission of Felix Loske.



A large body of data on this system was gathered by Felix Loske [36, 60], including quantitative measurements on cluster densities and sizes, as well as qualitative descriptions of cluster morphologies at and above room temperature. These show a transition from the completely filled two-layered triangular cluster at temperatures  $T > 315$  K to complex shapes with a branched structure and only a partly filled second layer at temperatures  $T < 300$  K and a temperature range of coexistence in between [Fig. 7(a-c)].

These observations together with the cluster density measurements within a temperature range of 96 K to 217 K (also including AFM images with visible cluster morphologies) form an abundance of data which we can try to reproduce with KMC simulations.

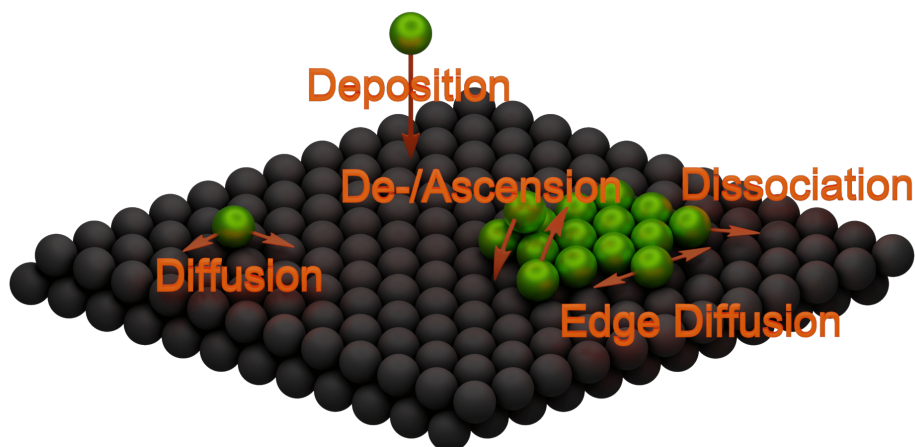
## II.2 EPITAXIAL CLUSTER GROWTH

In this section we are going to discuss the elementary transitions that govern the self-assembly of molecules on surfaces in deposition experiments, introduce different kinds of growth modes that can be observed and derive a model for the prediction of cluster densities from mean-field rate equation theory.

### II.2.1 ELEMENTARY DIFFUSIVE TRANSITIONS

The elementary transitions involved in the self-assembly of molecules on surfaces are mostly simple diffusion moves, but the variety of different possible transitions makes it hard to comprehend the importance of individual ones as well as the interplay between them. One quickly realizes that coming up with a model to accurately describe all the processes without an exploding amount of free parameters can be a difficult task. Examples of the most basic possible transitions are shown in Fig. 8. The deposition of particles can be done by MBE or vapour deposition and its rate can be varied in experiments, typically in the range of  $F = 0.01 - 1.0$  monolayers per minute (ML/min). The timespan of active particle deposition is called the deposition regime, the time after turning off the particle flux is called the post-deposition regime. During post-deposition, clusters can ripen into more favourable structures through diffusive transitions. This ripening process depends on the substrate temperature and can continue for several hours or even days, resulting in either an equilibrium or a kinetically trapped state. Throughout, evaporation of particles is generally possible but is negligible for most experiments [82, 83] and will not be considered in this work.

**Figure 8:** Basic elementary transitions during the cluster growth of molecules in deposition experiments.



Turning to the diffusive processes on the surface, they are assumed to be thermally activated, meaning that the transition rate  $k_i$  for a transition of type  $i$  can be described by an Arrhenius law [84]

$$k_i(T) = \nu_{0,i} \exp\left(-\frac{\Delta E_i}{k_B T}\right), \quad (\text{II.2})$$

providing the temperature dependence of those rates in terms of an energy barrier (or activation barrier)  $\Delta E_i$  and an attempt rate  $\nu_{0,i}$ .

For the case of free diffusion of single particles on the surface (FD), the energy barrier  $\Delta E_{FD}$  (or  $\Delta E_D$ ) is affected by the substrate-adparticle interaction on the first layer and by the adparticle-adparticle interaction for free diffusion on the second and higher layers ( $\Delta E_{FD,HL}$ ). It governs the mobility of the adparticle, thereby its ability to hit other adparticles or clusters to nucleate and ultimately it determines how many clusters form on the surface, as well as their size. If the diffusion process is composed of single jumps between neighbouring lattice sites, it can be characterized by a diffusion coefficient  $D$  that can be put into relation with  $\Delta E_{FD}$  and the attempt rate  $\nu_{0,FD}$  via

$$D = \frac{1}{4} \langle l^2 \rangle \nu_{0,FD} \exp\left(-\frac{\Delta E_{FD}}{k_B T}\right), \quad (\text{II.3})$$

with a mean squared jump length  $\langle l^2 \rangle$ . However, real diffusion processes often include jumps across multiple lattice sites or can even be hindered by impurities on the surface (so called impurity trapping [85]), which can result in an effective diffusion process between individual impurity sites. In such cases, one has to be careful with the application of Eq. II.3. If the diffusion coefficient  $D$  exhibits an Arrhenius-like behaviour [the data points  $(\log D, 1/T)$  form a line] in a given temperature range,  $D$  can be described by an effective attempt rate  $\nu_{0,FD,eff}$  and energy barrier  $\Delta E_{FD,eff}$ , which aren't necessarily in line with the parameters of the microscopic free diffusion transition ( $\Delta E_{FD}, \nu_{0,FD}$ ).

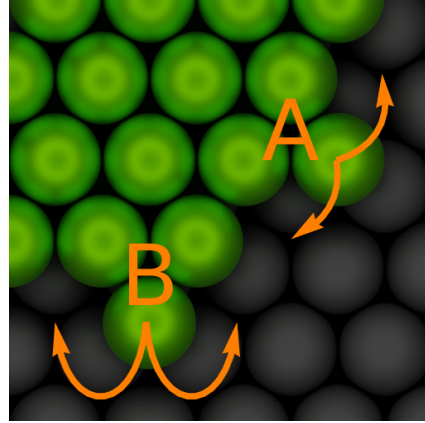
The remaining transition types in Fig. 8 are concerned with particles that are part of a cluster and as such affect the ordering and relaxation of the surface structures. These types of transitions typically depend on the configuration of the neighbouring particles and one can apply a bond counting approach for the energy barriers

$$\Delta E_i(n) = \Delta E_{i,0} + (n - \hat{n}_i) E_B \quad (\text{II.4})$$

where  $n$  is the initial number of bonds to neighbouring particles,  $\hat{n}_i$  is the number of bonds that can be sustained during the transition ( $\hat{n}_{ED-A/B} = 1$  for edge diffusion,  $\hat{n}_{Asc} = 2$  for ascension,  $\hat{n}_{Diss} = 0$  for dissociation),  $E_B$  is the bond energy and  $E_{i,0}$  is some base energy barrier for the transition  $i$ . As

an example, dissociation transitions  $i = \text{Diss}$  can be understood as normal free diffusion moves with added bond energies to the energy barrier, so for dissociation  $\Delta E_{\text{Diss},0} = \Delta E_{\text{FD}}$  is a valid assumption.

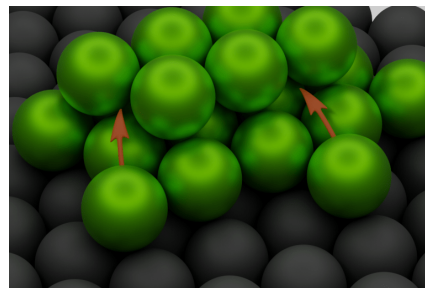
Edge diffusion is the transition by which clusters can relax their outer shape, e.g. to form rectangular, hexagonal or globular clusters. The energy barrier for edge diffusion can strongly depend on the adparticle-substrate interaction, e.g. resulting in different energy barriers for edge diffusion on edges with different alignment to the underlying substrate as shown in Fig. 9. In this example, the underlying substrate facilitates edge diffusion on the A edge while on the B edge it is largely hindered. Therefore it is necessary to distinguish two types of edge diffusion ( $i = \text{ED-A}/\text{ED-B}$ ).



**Figure 9:** Edge Diffusion along A and B Edges.

In this thesis, we will follow the convention that B edges are the ones with lower edge diffusion rates. Having strong differences between A and B edge diffusion rates can have a strong effect on the resulting cluster morphologies, creating triangular or even star-shaped islands as a result (as we will see in the next section). If edge diffusion is overall suppressed (e.g. at low temperatures) clusters grow into fractal shapes.

Ascension ( $i = \text{Asc}$ ) is the transition in which particles jump from one layer to the layer above. Its base energy barrier  $\Delta E_{\text{Asc},0}$  is governed by the binding energy to the substrate (or the adparticle layer below). If the adparticle has a strong interaction with the substrate (stronger than the adparticle-adparticle interaction), ascension will mostly be suppressed and clusters will grow as a monolayer (van der Merve growth) [83]. On the other hand, if the interaction with the substrate is very weak, clusters will grow as three-dimensional structures (Volmer-Weber growth). The  $\text{C}_{60}$  on  $\text{CaF}_2(111)$  system falls into an intermediate regime, where partly monolayered clusters can form structures on the second layer. If there are particles on the edge of the second layer, one may have to take into account the possibility of ascension to overhang sites (sites that are not fully supported by particles below) as shown in Fig. 10 which can even be facilitated by the attractive force of the particles on the second layer. In this figure, one can again see that a distinction between two types of overhang sites has to be made to be



**Figure 10:** Ascension to overhang sites.



accurate. The overall coordination of the two shown target sites, as well as the distance from the initial site to the target neighbours, differs on the two edges.

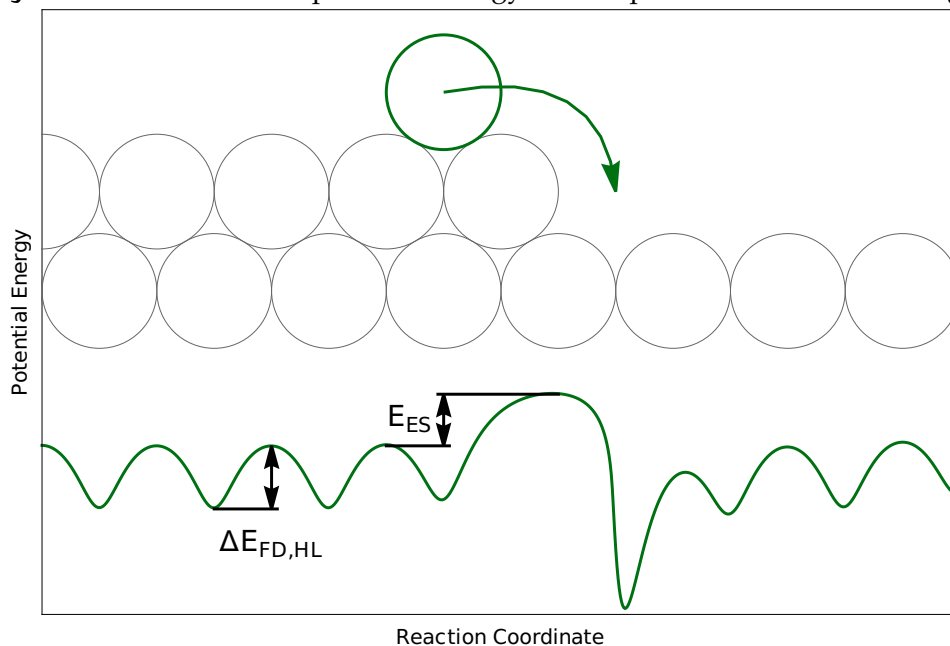
Finally, particles that are on top of an edge of a cluster can descend to the layer below ( $i = \text{Desc}$ ). For a descension move a slightly increased energy barrier in comparison to the free diffusion transitions has to be surmounted and the additional energy barrier is called the "Ehrlich-Schwöbel barrier"  $E_{ES}$

$$\Delta E_{\text{Desc}} = \Delta E_{\text{FD,HL}} + E_{ES}. \quad (\text{II.5})$$

A sketch of the potential energy landscape close to a cluster's edge is shown in Fig. 11. The magnitude of  $E_{ES}$  depends strongly on the range of the adparticle-adparticle interaction. While it can be very large for epitaxy of atoms, larger molecules like  $C_{60}$  have relatively weak Ehrlich-Schwöbel barriers. The interplay between deposition, diffusion, descension and ascension together govern the likelihood of nucleation and thereby formation of a new layer on top of a cluster.

After going through the details of all the individual elementary transitions, we have to conclude that a complete model for the epitaxy of molecules on surfaces is going to involve many energy barriers and attempt rates as free parameters, most of which are unavailable in the literature and not accessible in experiments. To get this problem under control, simplifications (like the bond counting approach II.4) are inevitable and computational methods have to be employed to determine values for the free parameters of the model.

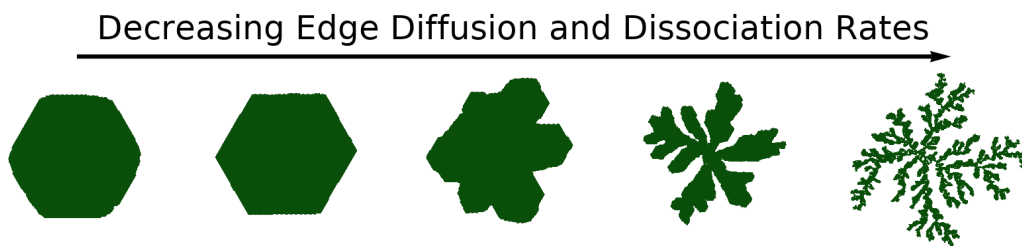
Figure 11: Sketch for the potential energy landscape close to a cluster's edge.



## II.2.2 CLUSTER SHAPES

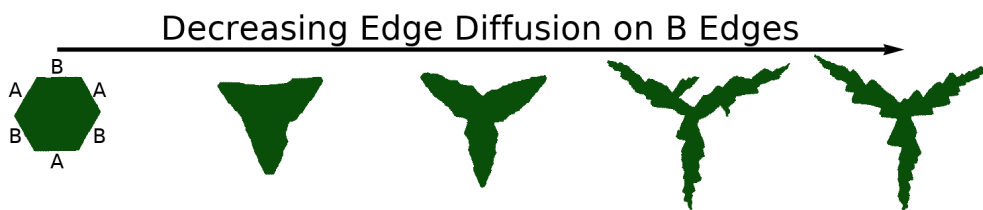
After introducing different kinds of elementary diffusive transitions, we now show the effect of certain transition types on cluster shapes to produce some of the commonly observed morphologies. We are also going to name those morphologies for later reference.

We start with the "hexagonal" clusters, which evolve when the edge diffusion and dissociation rates are high enough for the cluster to be able to relax into these compact shapes and if there is no major difference between the edge diffusion process on different edges of the cluster.



**Figure 12:** Change of cluster shapes upon decrease of edge diffusion and dissociation rates, lowering the clusters abilities to relax into compact shapes.

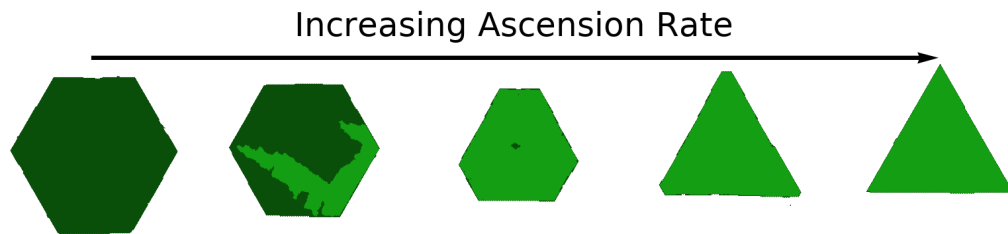
Fig. 12 shows how clusters evolve into shapes that we are going to call "fractal-dendritic" if edge diffusion and dissociation rates are suppressed. The lowering of these rates can be achieved by lowering the temperature of the system (which will also affect cluster sizes and densities) or by altering the bond strength  $E_B$  between adparticles (Eq. II.4).



**Figure 13:** Change of cluster shapes upon decrease of edge diffusion on edges of type B, making it harder for particles to find sites with high coordination on those edges.

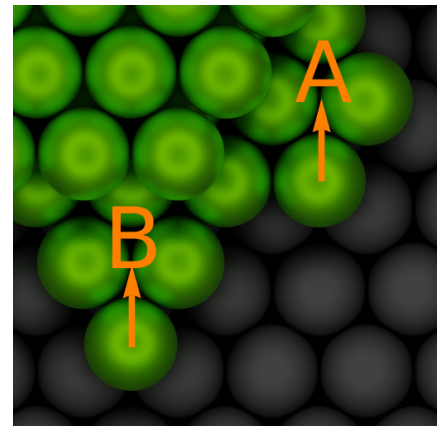
An effect that can be very strong in homoepitaxial growth is the difference between the edge diffusion rates on A and B edges (as illustrated in Fig. 9). The change in morphologies that arises when this difference is increased can be seen in Fig. 13. Slight differences can lead to "triangular" shapes, turning into "triangular star" shapes when the edge diffusion on B edges is strongly suppressed. The arms of these stars can also grow additional dendrites in the

directions of the A edges. The increased mobility on A step edges results in an increased ability to find stable positions (with three or four lateral neighbours) on these edges, making them grow faster than the B edges, on which the particles can behave more on the lines of hit-stick-dissociate. Moreover, whenever a particle is in a "corner" state where it can transition into the A and B step direction, it will prefer the A step edge, resulting in a net transfer of particles from B to A step edges.



**Figure 14:** Change of cluster shapes upon an increase of the ascension rate, making nucleation on the second layer more likely and increasing the speed of growth on higher layers. The second layer is coloured light green.

The effects of increasing the ascension rate are shown in Fig. 14. An obvious result is that starting with monolayer growth at low ascension rates, higher ascension rates lead to the growth of structures and eventually to full coverage of the second layer. Less intuitive is the observation of more and more "triangular" shapes of the two-layered clusters with a further increase of the ascension rate. The reason why these two-layered clusters tend to grow into triangles is illustrated in Fig. 15. The ascension of particles onto A edge positions requires only coordination of two adparticles on the edge of the cluster in common three-neighbour configurations while ascension onto B edge positions requires three adparticles in an unfavourable configuration (one particle being in a "kink" position, only being bound to the two adparticles upon which it can then ascend.). This makes it easier for clusters to grow into the directions of the A edge. It should be noted that in the ascension case, A and B edges refer to the edges of the second layer nucleus with respect to the first layer, not to the A/B edges of the first layer with respect to the substrate as was the case for the discussion of edge diffusion.



**Figure 15:** Ascension onto A and B edge positions.

### II.2.3 CLUSTER DENSITIES

In this section, we are going to derive an expression from nucleation theory for the prediction of cluster densities in deposition experiments based on experimental parameters. It is a slightly varied form of a derivation that can be found in Ref. [85, 86]. We start with the definition of rate equations for areal densities  $n_i$  of clusters composed of  $i$  adparticles,

$$\frac{dn_i}{dt} = \Gamma_{i-1} - \Gamma_i \quad i > 1, \quad (\text{II.6})$$

where  $\Gamma_i$  is defined as the net formation rate of size  $(i+1)$ -clusters from  $i$ -clusters. As a first simplification, we assume that clusters only grow by aggregation of single adparticles, based on a premise that clusters consisting of two or more adparticles are significantly less mobile. With this assumption, we can write  $\Gamma_i$  as

$$\Gamma_i = \sigma_i D n_1 n_i - \gamma_{i+1} n_{i+1}, \quad (\text{II.7})$$

where  $\gamma_i$  denotes the dissociation rate of adparticles from  $i$ -clusters,  $D$  the diffusion coefficient of single adparticles and  $\sigma_i$  the (dimensionless) capture number, quantifying the likelihood of an  $i$ -cluster to capture adparticles. So far, the set of equations II.6 is driven by the adparticle density  $n_1$  for which no rate equation was given. Assuming that evaporation of adparticles from the surface is negligible, and introducing the flux  $F$  of adparticles onto the surface per surface area, we can complete the set of rate equations by writing down the rate equation for single adparticles as

$$\frac{dn_1}{dt} = F - 2\Gamma_1 - \sum_{i \geq 2} \Gamma_i, \quad (\text{II.8})$$

However, the large amount of unknown free parameters in this set of equations makes further simplifications necessary. We start by assuming the existence of a critical cluster size  $i^*$  above which clusters become stable, i.e. the dissociation rate  $\gamma_i$  from clusters with sizes  $i > i^*$  vanishes ( $\gamma_i = 0$ ). In practice, this critical cluster size  $i^*$  can depend on the type of adparticle/substrate, on the flux  $F$  and the temperature  $T$ . To be able to arrive at a result for the overall density of stable clusters  $N$ , we define it as

$$N = \sum_{i > i^*} n_i. \quad (\text{II.9})$$

Assuming that the net cluster formation rate  $\Gamma_i$  vanishes for  $i \rightarrow \infty$ , the differential equation for  $N$  takes the compact form

$$\frac{dN}{dt} = \sum_{i>i^*} \Gamma_{i-1} - \Gamma_i = \Gamma_{i^*} = \sigma_{i^*} D n_1 n_{i^*} \quad (\text{II.10})$$

To handle the unstable clusters ( $i \leq i^*$ ), we assume that they are in thermal equilibrium, meaning that the net formation rates  $\Gamma_i$  vanish for those cases. With this assumption, we can approximate the probability of finding a cluster of size  $i$  on an adsorption site with area  $\Omega$  with the Walton relation [87]

$$\Omega n_i \approx (\Omega n_1)^i e^{E_i/k_B T}. \quad (\text{II.11})$$

Eq. II.11 is composed of a product between a Boltzmann factor, incorporating the energy  $E_i$  that is gained by forming a cluster of size  $i$ , and the probability of finding  $i$  independent adparticles on a lattice site of size  $\Omega$ ,  $(\Omega n_1)^i$ . Inserting Eq. II.11 into Eq. II.10 yields

$$\frac{dN}{dt} = \sigma_{i^*} D n_1^{i^*+1} \Omega^{i^*-1} e^{E_{i^*}/k_B T}. \quad (\text{II.12})$$

Finally, we introduce a definition for the average capture number  $\bar{\sigma}$

$$\bar{\sigma} = \frac{1}{N} \sum_{i>i^*} n_i \sigma_i, \quad (\text{II.13})$$

to simplify Eq. II.8 into

$$\frac{dn_1}{dt} = F - \frac{dN}{dt} - \bar{\sigma} D n_1 N \quad (\text{II.14})$$

and obtain a closed set of equations for the adparticle and cluster densities with Eqs. II.14 and II.12. Although the remaining parameters  $\bar{\sigma}$  and  $\sigma_i$  can vary with  $N/i/n_1$  we assume that they are constants for the following discussion of solutions to the equations. A detailed discussion on these parameters can be found in Refs. [83, 85].

The first solution we can give is valid in the so-called transient nucleation regime, where  $N \ll n_1$ . In this early stage of the deposition process, the loss terms to the adparticle densities in Eq. II.14 can be ignored, leading to a linear increase of  $n_1$  with the total coverage  $\Theta$

$$n_1 \sim \Theta = \Omega F t, \quad (\text{II.15})$$

resulting in a cluster formation rate that increases according to

$$\frac{dN}{dt} \sim \Theta^{i^*+1}. \quad (\text{II.16})$$

Secondly, we can make statements about the steady state nucleation regime, where the cluster formation rate can be neglected relative to the aggregation of adparticles to clusters and the adparticle density is in a steady state:

$$\frac{dN}{dt} \ll \bar{\sigma} D n_1 N, \quad \frac{dn_1}{dt} \approx 0. \quad (\text{II.17})$$

Putting the assumptions [II.17](#) into [Eq. II.14](#), we obtain

$$n_1 \approx \frac{F}{\bar{\sigma} D N} \quad (\text{II.18})$$

for the adparticle density. Going a step further, we can now derive a formula for the cluster density per adsorption site  $\Omega N$  by first inserting [Eq. II.18](#) into [Eq. II.12](#),

$$\frac{dN}{dt} = \sigma_{i^*} D \left( \frac{F}{\bar{\sigma} D N} \right)^{i^*+1} \Omega^{i^*-1} e^{E_{i^*}/k_B T},$$

and then integrating via separation of variables:

$$\begin{aligned} N^{i^*+1} dN &= \sigma_{i^*} D \left( \frac{F}{\bar{\sigma} D} \right)^{i^*+1} \Omega^{i^*-1} e^{E_{i^*}/k_B T} dt \\ \frac{1}{i^*+2} N^{i^*+2} &= \sigma_{i^*} D^{-i^*} \left( \frac{F}{\bar{\sigma}} \right)^{i^*+1} \Omega^{i^*-1} e^{E_{i^*}/k_B T} t, \end{aligned} \quad (\text{II.19})$$

which after solving for  $\Omega N$  and some restructuring leads us to our final result for the cluster density per adsorption site:

$$\Omega N \approx \eta(\Theta, i^*) \left( \frac{\Omega^2 F}{D} \right)^{\frac{i^*}{i^*+2}} e^{E_{i^*}/(i^*+2)k_B T}, \quad (\text{II.20})$$

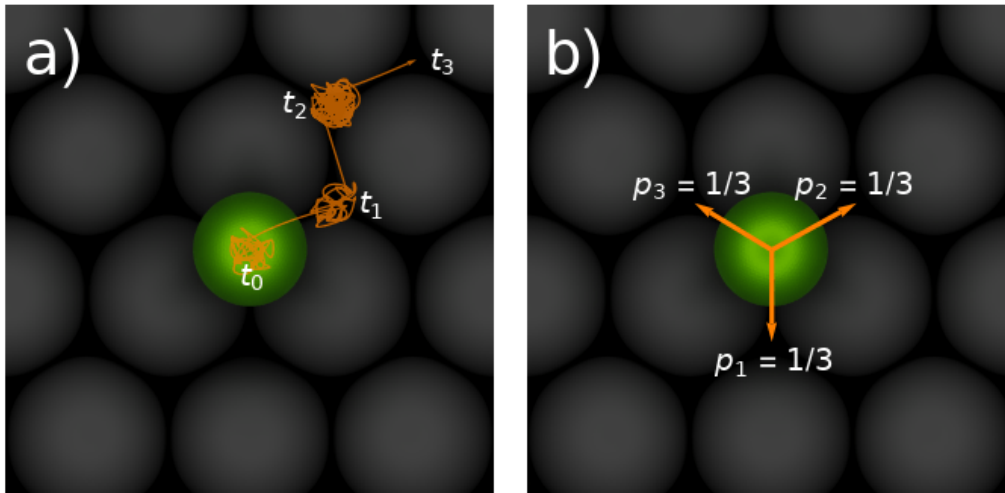
where the definition  $\eta(\Theta, i^*) = [(\Theta \sigma_{i^*} / \bar{\sigma}^{i^*-1})^{1/(i^*+2)}]$  has been used. [Eq. II.20](#) is a central result of nucleation theory as it can be used to gain insight into the physical parameters of the underlying diffusion and assembly process (like  $D$ ,  $\Delta E_D$ ,  $\nu_{0,D}$ ,  $i^*$  and  $E_{i^*}$ ) only by looking at the resulting macroscopic cluster densities of an experiment [[36](#), [88](#)].

## II.3 KINETIC MONTE CARLO

Algorithms that make use of (pseudo)random number generators [App. B.1] to solve problems are referred to as "Monte Carlo" (MC) methods. In computer simulations of physical systems, MC methods have now been in use for the better part of a century. One prominent and successful example is the metropolis algorithm [89], which can be used to sample states of a system around energy minima by applying small random changes to the system and accepting or rejecting them based on a criterion (Metropolis criterion) that prefers changes which reduce the overall potential energy. The Metropolis algorithm can be very useful to learn about equilibrium properties of a system, however it can not reproduce the dynamic evolution of a system, as the changes one applies to the system do not represent physical trajectories. Systems that evolve into kinetically trapped out-of-equilibrium states, as is often the case in epitaxy experiments, are therefore out of scope of such simulation methods. A class of methods that was developed to handle certain types of such systems (so-called "infrequent-event systems") are "kinetic Monte Carlo" (KMC) methods (e.g. the Gillespie algorithm [50, 51]). These methods can be very efficient in simulating the dynamic evolution of a system, e.g. they can be able to calculate trajectories of epitaxial growth at time scales of hours and length scales of micrometres, making it a natural choice for our system of interest. In this section we are going to introduce the basic principles of KMC methods. As main sources we are using Refs. [49, 90] and my master's thesis [86].

### II.3.1 INFREQUENT-EVENT SYSTEMS

"Infrequent-event systems" are systems that spend a large amount of time in metastable, discretized states with rare relatively quick transitions between the individual states. A prime example for such systems is the decay of radioactive atoms, which can spend time scales of years in a metastable state to then finally decay into a new state in a matter of femtoseconds. A different example is the surface diffusion processes that we are interested in. The problem that such systems pose to conventional molecular dynamics simulations is that if one wants to simulate a diffusion trajectory, one has to spend a lot of CPU time to calculate the microscopic oscillations of the diffusing particle around the individual lattice sites [Fig. 16a], leading to a very slow progression of the macroscopic diffusion process. Although such trajectories from MD simulations are basically deterministic, if you are not looking at the fine details and generate many trajectories with different initial parameters, the whole process can be understood as a random walk with randomly distributed waiting times between the jumps. This is the understanding on which KMC simulations are



**Figure 16:** Surface diffusion as an example for an infrequent-event system. a) An example trajectory as it could be the result of molecular dynamics simulations. The major part of the trajectory is spent oscillating around the individual lattice sites. b) The result of such a trajectory can as well be modelled by random transitions into one of the three possible directions with equal probability.

based. Instead of calculating the complete trajectories, one instead generates a random walk by randomly choosing one of the possible diffusion directions with equal probability. The time evolution of the system can be handled if the distribution of waiting times between transitions is known. Like in many physical systems, we can assume that the probability to transition from one state to another per time interval stays constant, which leads to an exponential distribution of the waiting times

$$p(t_{\text{wait}}) = ke^{-kt}, \quad (\text{II.21})$$

with an overall transition rate  $k$ , which can be calculated as a sum of individual transition rates ( $k = k_1 + k_2 + k_3$  for our example of a particle with three possible diffusion directions). To advance the time in a KMC simulation, we have to draw a random number according to the exponential distribution (Eq. II.21). This can be achieved by drawing a uniformly distributed random number  $r \in (0, 1)$  and transforming it into an exponentially distributed waiting time via

$$t_{\text{wait}} = -\frac{1}{k} \ln(r) \quad (\text{II.22})$$

If the transition rates  $k_i$  are chosen correctly (as they would be measured in the corresponding MD simulations) the overall macroscopic diffusion processes resulting from such randomly generated trajectories are going to be indistinguishable from the actual trajectories in MD simulations.



While we have only discussed a very simple example so far, which probably could be simulated without any sophisticated algorithm, the introduction of multiple particles with interactions between them will quickly make the situation more complicated. In the next section, we are going to go into more detail on a specific KMC algorithm that can handle such systems, the "rejection-free KMC" algorithm.

### II.3.2 REJECTION-FREE KMC

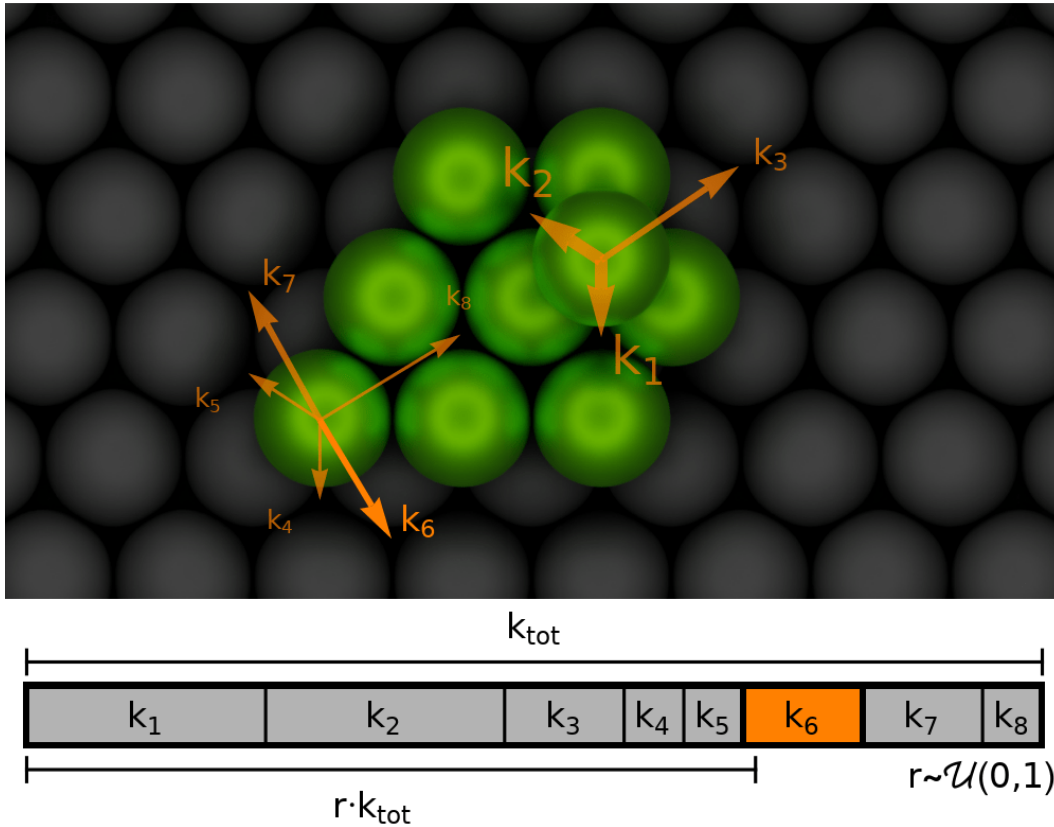
The algorithm we have chosen to use in our implementation is the rejection-free kinetic Monte Carlo (rfKMC) method. It is called "rejection-free" because every step in the simulation applies some change to the system, which is in contrast to some other MC methods (like the Metropolis algorithm) where changes can be rejected. This can be an obvious advantage, as no computation time is being wasted on changeless steps. It is sometimes also referred to as the "residence-time" algorithm or "BKL" algorithm, referring to Bortz, Kalos and Lebowitz, who described the algorithm in their 1975 paper [91].

To describe the algorithm, suppose we have a system that has  $N$  different possible transitions it can go through and we assume that we know the corresponding transition rates  $\{k_1, \dots, k_N\}$ . To decide which of those transitions is next to execute, we represent them as intervals, the lengths of which are proportional to the values of their transition rates  $k_i$ . Putting these intervals next to each other, we obtain a total interval representing the total transition rate  $k_{\text{tot}} = \sum_{i=1}^N k_i$ . The pathway selection is now done by drawing a uniformly distributed random number  $r \sim \mathcal{U}(0, 1)$  and multiplying it by  $k_{\text{tot}}$  to obtain a random position in this concatenation of intervals (See Fig. 17). This random position falls into one of the intervals of the transitions, thereby choosing it as the pathway to execute in this step of the simulation. The probability of choosing transition  $i$  with this algorithm is proportional to the transition rate  $k_i$  and given via  $k_i/k_{\text{tot}}$ , just as required for independent transitions with exponentially distributed waiting times (See appendix B.3). To write this procedure into an algorithm, we construct an array of partial sums

$$s_j = \sum_{i=1}^j k_i \quad j \in \{1, 2, \dots, N\}. \quad (\text{II.23})$$

and look for the smallest  $j$  which satisfies  $s_j > rk_{\text{tot}}$  to decide which transition is going to be executed.

To advance the time after executing a transition we proceed as described earlier, by drawing an exponentially distributed waiting time according to Eq. II.21 with  $k = k_{\text{tot}}$ . Afterwards, the possible transitions and the corresponding transition



**Figure 17:** Illustration of the pathway selection in rfKMC simulations. In this example we have two mobile particles, particles with more than three neighbours are frozen. The possible transitions are shown with orange arrows, the thickness of which represent the magnitude of the corresponding transition rate. As described in the main text, the drawing of a random number  $r \sim \mathcal{U}(0,1)$  here decides that transition 6 is the one that gets executed, moving the particle to the bottom left corner of the cluster via edge diffusion.

rates have to be determined for the new state of the system before executing the next step of the simulation.

To sum up this section, we will put the discussed procedure into a well-defined algorithm. We define  $\mathcal{S} = \{1, 2, 3, \dots\}$  as the set of indices describing the possible microstates of the system. The index  $i \in \mathcal{S}$  describes the current state of the system and transition rates from this state into other states  $j$  are denoted as  $k_{ij}$ . The transitions  $i \rightarrow j$  should be categorizable into discrete types of transitions (e.g. diffusion/edge diffusion/ascension/... with  $n \in \{1, 2, \dots\}$  initial neighbours) for which the corresponding transition rates should be known in advance through some kind of rate model. We start with the system being in any state  $i$  at time  $t = 0$  and loop over the following procedure:

1. Determine which transitions  $i \rightarrow j$  are possible and assign the transition rates  $k_{ij}$  according to the given rate model. Impossible transitions are assigned  $k_{ij} = 0$ .

2. Calculate the total transition rate  $k_{\text{tot}}$  as well as the partial sums  $s_n$

$$k_{\text{tot}} = \sum_{j \in \mathcal{S}, j \neq i} k_{ij} \quad s_n = \sum_{j=1}^n k_{ij}$$

3. Draw a random number from the uniform distribution  $r \sim \mathcal{U}(0,1)$  and determine which transition is executed next by finding the index  $n$  that satisfies

$$s_{n-1} < rk_{\text{tot}} \leq s_n. \quad (\text{II.24})$$

4. Execute transition  $i \rightarrow n$ .
5. Draw another random number from the uniform distribution  $r_t \sim \mathcal{U}(0,1)$  and update the time of the system via

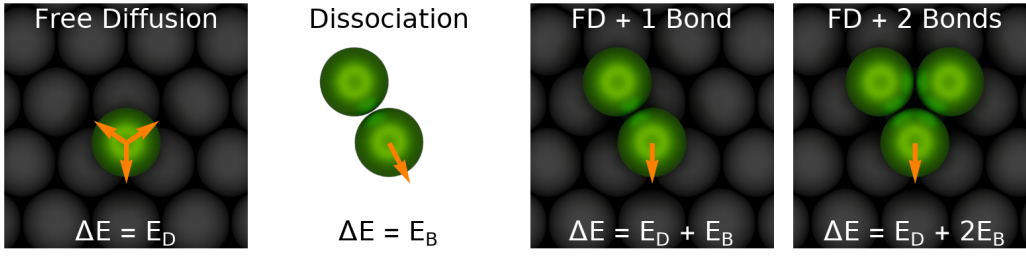
$$t \rightarrow t + \frac{1}{k_{\text{tot}}} \ln \left( \frac{1}{r_t} \right)$$

6. Go back to step 1.

Steps 1 to 3 are where the main computational effort is going to be spent for any system of decent size. Coming up with a data structure and algorithms that enable fast refreshing of the transition rates and partial sums after each step as well as fast searching for the transition that satisfies Eq. II.24 is crucial when implementing an rfKMC simulation. We provide details on our implementation in the appendix chapter A.

### II.3.3 TRANSITION RATE MODELLING

If one wants to reproduce a physical system like the epitaxial growth of  $C_{60}$  in KMC simulations, it is crucial to come up with an accurate model for the elementary transition rates of the system. As mentioned earlier, the assumption of an Arrhenius law (Eq. II.2) will give us the temperature dependence of the individual transition rates, however, this will still leave energy barriers and attempt rates as unknown parameters that have to be determined for every transition type. It is infeasible to determine all of these parameters experimentally and it would even be a huge task to do it computationally, so some simplifications have to be applied to reduce the amount of parameters. A very important simplification is the application of a bond counting approach, which is based on the assumption of additive interactions (illustrated in Fig. 18). Because the interaction range of  $C_{60}$  molecules with each other is relatively



**Figure 18:** Application of additive interactions to energy barriers. from the known energy barriers of free diffusion on a substrate and dissociation of a dimer off the substrate, one can deduce the energy barriers for dissociation of a particle from multiple neighbours on a substrate.

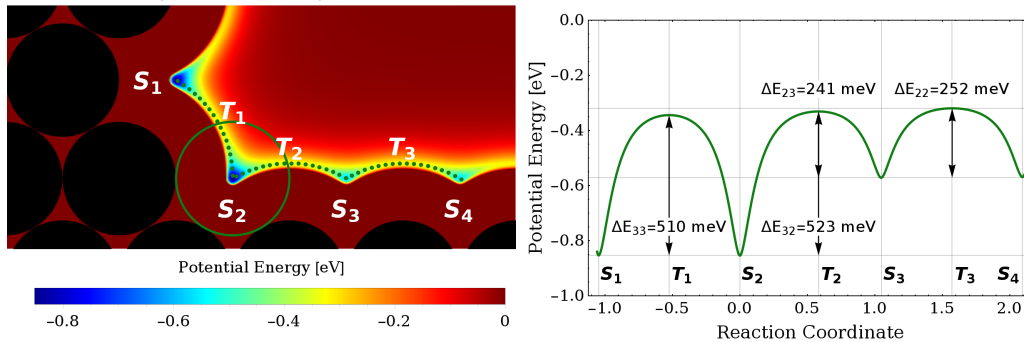
short-ranged, we can even assume that the energy barriers only depend on the lateral neighbours of the initial site, neglecting the neighbourhood of the target sites (See Fig. 19). This leads to a formula for the transition rate that depends on the transition type  $i$  ( $i \in \{\text{FD/Diss, ED-A/B, Asc, Desc}\}$ ), the number of initial neighbours  $n$  and the temperature  $T$ ,

$$k_i(n, T) = \nu_{0,i}(n) \exp\left(-\frac{\Delta E_i(n)}{k_B T}\right), \quad (\text{II.25})$$

where the energy barriers can be described by the aforementioned formula,

$$\Delta E_i(n) = \Delta E_{i,0} + (n - \hat{n}_i)E_B, \quad (\text{II.26})$$

**Figure 19:** Left: Density plot of the potential energy of a sample  $C_{60}$  molecule (dark green circle) in a given configuration of fixed cluster molecules (black disks, viewed from above) interacting via the Girifalco potential (the substrate is omitted). The dotted dark green line represents a possible reaction path connecting the four states  $S_1, \dots, S_4$  passing through the transition states  $T_1, T_2$  and  $T_3$ . Right: Plot of the potential energy along the dotted reaction path. The energy barriers  $\Delta E_{ij}$  ( $i$  the number of initial neighbours,  $j$  the number of final neighbours) are only weakly affected by  $j$ . Reprinted with permission from Ref. [1]. Copyright 2020 by the American Physical Society.



reducing the energy parameters to the bond energy  $E_B$  and the base energy barriers  $\Delta E_{i,0}$ .

This leaves us with the attempt rates  $\nu_{0,i}(\mathbf{n})$  for which no established modelling approaches exist. Leaving them all as free parameters would give the overall model too many unnecessary degrees of freedom, considering that the impact of having the exact attempt rates is relatively small in comparison to the energy barrier parameters. Therefore, the attempt rates are often assumed to have a single constant value

$$\nu_{0,i}(\mathbf{n}) = \nu_0 \quad (\text{II.27})$$

or sometimes (preferably for atomistic systems) they are even calculated with a dependence on the temperature [55, 92–94] via

$$\nu_{0,i}(\mathbf{n}) = \frac{2k_B T}{h}, \quad (\text{II.28})$$

with the Planck constant  $h$ . Alternatively, to leave a little more freedom in these parameters, one can add constraints via the detailed balance condition

$$\frac{k_i(\mathbf{n}, T)}{k_i(\mathbf{m}, T)} \stackrel{!}{=} \frac{k_j(\mathbf{n}, T)}{k_j(\mathbf{m}, T)}, \quad \xrightarrow{\text{II.25, II.26}} \quad \frac{\nu_{0,i}(\mathbf{n})}{\nu_{0,i}(\mathbf{m})} = \frac{\nu_{0,j}(\mathbf{n})}{\nu_{0,j}(\mathbf{m})} := c_{nm} \quad (\text{II.29})$$

leading to constant ratios  $c_{nm}$  between attempt rates for different numbers of initial neighbours that are independent of the transition type  $i$ . This effectively introduces the parameters  $c_{12}, c_{23}, c_{34}, c_{45}, c_{56}$  but reduces the free attempt rates per transition type to just a single one,  $\nu_{0,i}(\mathbf{n}_0)$ , from which the others can be calculated. To restrict this modelling a little more, one can decide to set the ratios  $c_{nm}$  to one between larger numbers of neighbours (e.g.  $c_{34} = c_{45} = c_{56} = 1$ ), as those transition rates are usually strongly suppressed anyway and have limited impact on the result of the simulations.

Depending on how many of these discussed simplifications one chooses to use, the resulting model will still have around ten or more free parameters which one has to somehow determine. To enable the fitting of these parameters, in the following subsections we are going to discuss two possible computational approaches to obtain data on transition rates, the harmonic transition state theory (HTST) and MD simulations.

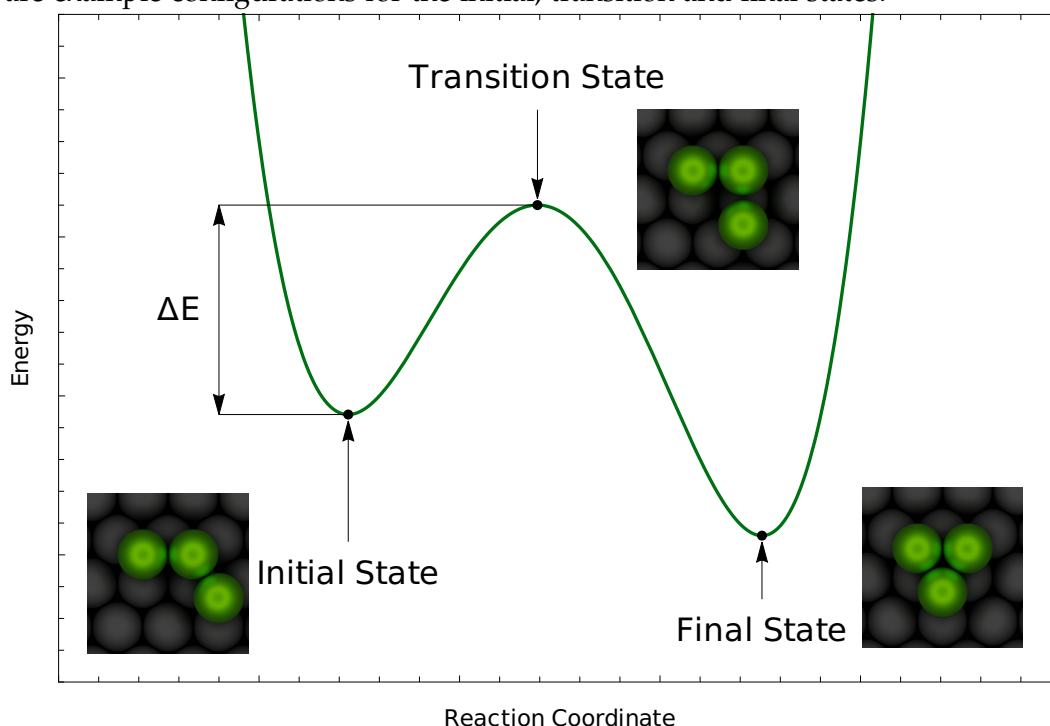
### II.3.4 HARMONIC TRANSITION STATE THEORY

Transition state theory (TST) methods have been developed for the approximation of rate constants of chemical reactions in the early 20th century [95–97]. In TSTs a transition is characterized by three states, an initial, a transition and a final state. The initial and final states are nearby local minima of the potential energy surface (PES) and they are connected via a saddle point in the PES, the transition state. An energetically optimal pathway that connects these three states is called a "minimum energy path" (MEP) or "reaction path". A parameter that describes a parametrization of an MEP is called a "reaction coordinate". Harmonic transition state theory (HTST), also known as vineyard theory [98], is a TST method that can approximate the activation barriers  $\Delta E$  and attempt rates  $\nu_0$  (as used in the Arrhenius law) for a transition from the PES close to its initial and transition state. The resulting HTST transition rate can be calculated via

$$k^{\text{HTST}} = \frac{\prod_i^N \nu_i^{\text{init}}}{\prod_i^{N-1} \nu_i^{\text{trans}}} \exp\left(-\frac{\Delta E}{k_B T}\right), \quad (\text{II.30})$$

where the energy barrier  $\Delta E$  is the difference in potential energy between the initial and transition state [Fig. 20] and  $\nu_i^{\text{init}}, \nu_i^{\text{trans}}$  are the normal mode frequencies at the initial and transition state respectively. While the product

**Figure 20:** Illustration of the energy landscape along a reaction path. The insets are example configurations for the initial, transition and final states.



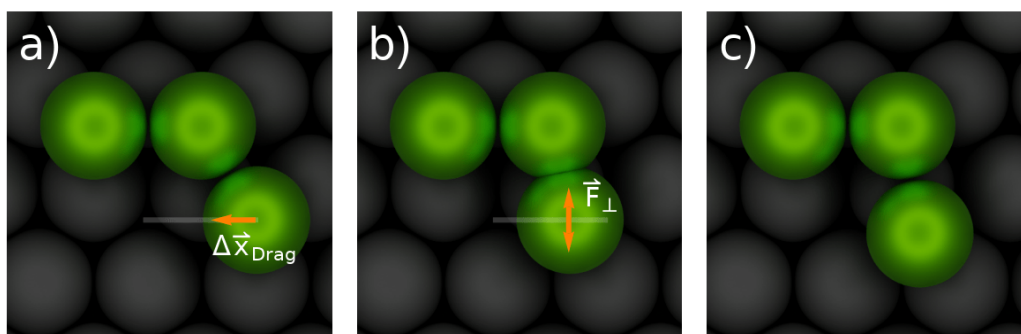
in the numerator of the fraction goes over all  $N$  normal mode frequencies of the initial state, the product in the denominator only contains the  $N - 1$  normal mode frequencies that are perpendicular to the reaction path. It has been shown that the approximated transition rates from HTST can come very close to the actual transition rates of a system [99, 100].

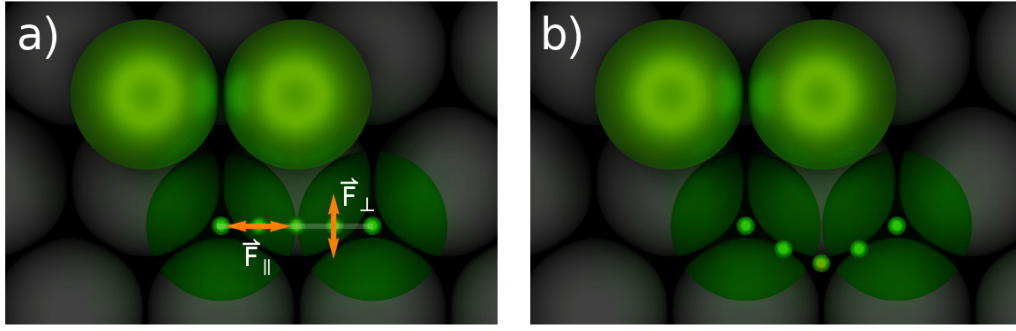
### *Minimum Energy Paths*

Methods to determine minimum energy paths are often used in TST as a way to find the transition state. Here, we are going to discuss two approaches for MEP calculation (based on Ref. [101]) which also find use in this thesis, the "drag" and the "nudged elastic band" (NEB) method [102–104].

A very simple and intuitive method for finding an MEP is the drag method. Starting from the initial state, one chooses one of the  $N$  degrees of freedom of the system to be the "drag coordinate" which is incremented in small steps to drive the system from the initial to the final state [Fig 21a]. After each increment, the  $N - 1$  degrees of freedom perpendicular to the drag coordinate are relaxed to minimize the potential energy  $E$  of the system [Fig 21b], shifting it towards a point on the MEP [Fig 21c]. The point of maximum potential energy achieved during the drag is considered the transition state. While this method is conceptually very simple, its successful application mostly depends on the arbitrary choice of the drag coordinate and a bad choice can result in an MEP that never reaches the actual vicinity of the transition state. The drag method can certainly still be applied to many cases, but one has to be aware that there are examples where it does not yield good results [105, 106] and the

**Figure 21:** Simple two-dimensional illustration of an iteration of the drag method for a particle diffusing from a one neighbour state to a two neighbour state in a fixed environment (only the diffusing particles coordinates are varied). a) A direct line connecting the initial and final state is chosen as the drag coordinate and the particle is displaced in this direction. b) After displacement, the particles coordinates are relaxed perpendicular to the drag coordinate to minimize the potential energy. c) The particle has relaxed into a position on the MEP and can be dragged further in the next iteration.





**Figure 22:** Simple two-dimensional illustration of the NEB method. a) On an initial path that connects the initial with the final state, “beads” are placed equidistantly as representations of the systems state at several points along the MEP. b) These beads are “nudged” as described in the main text to obtain the final MEP.

simultaneous application of other supporting methods (like the NEB method) may be well advised.

In the NEB method, a string of interconnected images  $\{\mathbf{S}_0, \mathbf{S}_1, \dots, \mathbf{S}_N\}$  of the system is created (also referred to as the “elastic band”) to form a discrete initial path connecting the initial state  $\mathbf{S}_0$  with the final state  $\mathbf{S}_N$  [Fig. 22a]. Here, the states are represented by coordinate vectors  $\mathbf{S}_i = (x_1, x_2, \dots, x_D)$ , where  $D$  is the dimensionality of the system. While  $\mathbf{S}_0$  and  $\mathbf{S}_N$  are fixed, the other  $N - 1$  elements of the string are relaxed into the MEP through so-called “nudging” by the application of two forces to each of the images,

$$\mathbf{F}_i = \mathbf{F}_{i,\parallel} + \mathbf{F}_{i,\perp}. \quad (\text{II.31})$$

The perpendicular force  $\mathbf{F}_{i,\perp}$  pushes the images towards points on the MEP by pushing them in the direction of the projected energy gradient

$$\mathbf{F}_{i,\perp} = -\nabla E(\mathbf{S}_i) + \nabla E(\mathbf{S}_i) \cdot \hat{\tau}_i \hat{\tau}_i, \quad (\text{II.32})$$

where  $\nabla$  is the “Nabla” operator  $\nabla = (\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_D})$  and the  $\hat{\tau}_i$  are estimated tangents to the elastic band that can e.g. be calculated via

$$\tau'_i = \frac{\mathbf{S}_i - \mathbf{S}_{i-1}}{|\mathbf{S}_i - \mathbf{S}_{i-1}|} + \frac{\mathbf{S}_{i+1} - \mathbf{S}_i}{|\mathbf{S}_{i+1} - \mathbf{S}_i|}, \quad \hat{\tau}_i = \frac{\tau'_i}{|\tau'_i|}. \quad (\text{II.33})$$

A more detailed discussion on the estimation of the tangents  $\hat{\tau}_i$  (including a better estimate) can be found in Ref. [101]. The second term of Eq. II.32 removes the components of the energy gradient that are parallel to the elastic bands tangent. The parallel force  $\mathbf{F}_{i,\parallel}$  is a spring force that makes the band of



images "elastic" by ensuring that even spacing between the individual states is kept during the nudging. It can be calculated via

$$\mathbf{F}_{i,\parallel} = \kappa(|\mathbf{S}_{i+1} - \mathbf{S}_i| - |\mathbf{S}_i - \mathbf{S}_{i-1}|)\hat{\tau}_i, \quad (\text{II.34})$$

with a spring constant  $\kappa$ . Iterative displacement of the elastic band in small increments via the nudging forces  $\mathbf{F}_i$  leads to a convergence of the states  $\mathbf{S}_i$  onto the MEP and to a convergence of the nudging forces  $\mathbf{F}_i$  to zero. Since its first formulation, the NEB method quickly became a popular solution for the determination of MEPs of several diffusion processes [107–111] and has also been improved upon, e.g. with the climbing image (CI-NEB) modification [112] which further improves the methods ability to find the exact transition state. While the NEB method also requires an arbitrary guess for the initial path, a simple linear interpolation between  $\mathbf{S}_0$  and  $\mathbf{S}_N$  is a good choice in most cases. If there are multiple MEPs present, some sort of sampling of the various MEPs needs to be carried out to be able to find the optimal one.

### II.3.5 TRANSITION RATE MEASUREMENTS IN MD SIMULATIONS

The approach that we are using the most in this thesis for the determination of transition rates is to measure them directly in molecular dynamics simulations. The details behind the molecular dynamics method are going to be discussed in the next section, here we only explain what we are measuring in the MD simulations and how we derive energy barriers and attempt rates from those measurements.

One starts with the setup of configurations in which the transitions of interest can be measured and are unlikely to be disturbed by other transitions. An example is shown in Fig. 23 (left) for transitions of a particle from a two neighbour initial state, where the other particles are all in a four neighbour state and therefore are unlikely to leave their lattice site. MD simulations are then run in a range of temperatures in which the transitions of interest are likely to occur (for the example given in Fig. 23, this range is  $T \in (500, 700)\text{K}$ ). Each MD simulation only runs as long as the particle stays in its initial state. As soon as it transitions away, the residence time  $\tau$  in the initial state and the type of occurred transition  $i$  are written out and the MD simulation is restarted with slightly varied initial conditions. This is repeated until enough trajectories have been simulated to capture data on all transitions of interest.

As a result for each temperature  $T$ , one gathers a set of residence times  $\{\tau_1, \tau_2, \dots, \tau_{N(T)}\}$ , where  $N(T)$  is the total number of observed trajectories, and counts of the observed transitions  $N_i(T)$  ( $i \in \{\text{ED-A, Diss, Asc}\}$  for the exam-

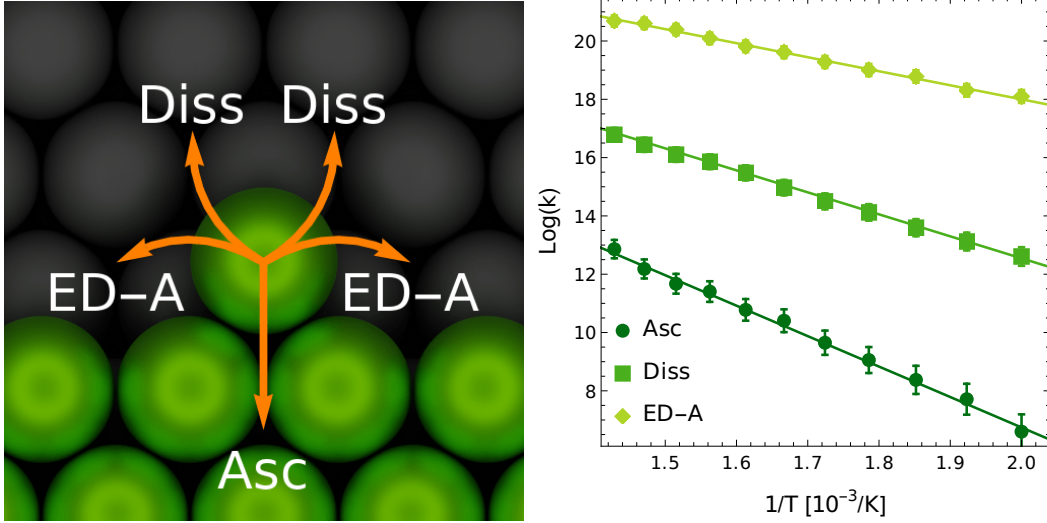


Figure 23: Left: Example configuration for the determination of transition rates from a two neighbour initial state in MD simulations. Right: Derived Arrhenius plot including fits for the extraction of energy barriers and attempt rates.

ple of Fig. 23). We are presenting two ways to derive the total transition rate  $k_{\text{tot}}(T)$  out of the initial state from those results. The first is a direct calculation with the (unbiased) estimator for the exponential distribution

$$\hat{k}_{\text{tot}}(T) = \frac{N(T) - 1}{t_{\text{tot}}(T)} \quad t_{\text{tot}}(T) = \sum_{j=1}^{N(T)} \tau_j. \quad (\text{II.35})$$

If  $N$  is large, it may be more convenient to use the (biased) maximum likelihood estimator  $\hat{k}'_{\text{tot}} = N/t_{\text{tot}}$ . The second way simultaneously acts as a double check to see if the residence times are indeed exponentially distributed (as assumed by Eq. II.35 and by KMC algorithms). If they are exponentially distributed, one can derive a total transition rate  $k_{\text{tot}}(T)$  from the distribution of residence times by fitting an exponential fit

$$p(\tau) = k_{\text{tot}}(T)e^{-k_{\text{tot}}(T)\tau} \quad (\text{II.36})$$

to the histogram of residence times [Fig. 24]. From the  $k_{\text{tot}}$  that is either derived from Eq. II.35 or from the fit II.36, one can then estimate the transition rates for the individual transition types via

$$\hat{k}_i(T) = \frac{N_i(T)}{N(T)}k_{\text{tot}}(T). \quad (\text{II.37})$$

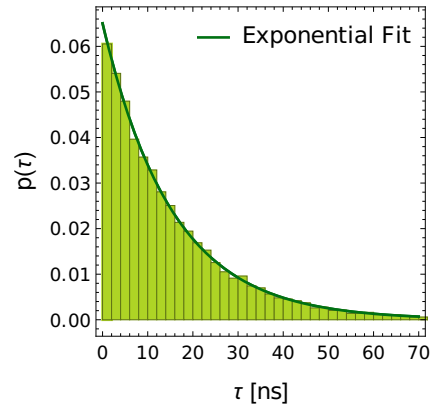


Figure 24: Example probability distribution of residence times.

Finally, the estimates for the transition rates that have been gathered in a range of temperatures can be used to derive energy barriers and attempt rates from so-called Arrhenius plots. By applying a logarithm to the Arrhenius law (Eq. II.2) we obtain the equation

$$\log(k_i(T)) = \log(\nu_{0,i}) - \frac{\Delta E_i}{k_B T}. \quad (\text{II.38})$$

By plotting the logarithmic transition rates  $\log(k_i(T))$  against the inverse temperature  $1/T$ , we can verify that the rates indeed behave as described by the Arrhenius law if they fall on a line [Fig. 23 right]. Energy barriers  $\Delta E_i$  and attempt rates  $\nu_{0,i}$  can then be extracted from the fit parameters of a linear fit of the form  $y(x) = mx + b$  to the data points  $(1/T, \log(k_i(T)))$  via

$$\nu_{0,i} = e^b \quad \Delta E_i = -\frac{m}{k_B}. \quad (\text{II.39})$$

While this whole method of determining transition rates from MD simulations can overall take more computational effort than TST methods, it has the clear advantage of needing fewer assumptions. One only has to assume the validity of some interaction potentials for the substrate atoms and adparticles with one another. The validity of the Arrhenius law and the exponential distribution of waiting times can then be verified by the results of the simulation, which adds justification to the use of the derived transition rate parameters ( $\Delta E_i, \nu_{0,i}$ ) in KMC simulation models.

### *Free Diffusion and the Mean Squared Displacement*

A special case in the determination of transition rates is the free diffusion process. Here, we are not necessarily interested in the exact escape rate from a certain lattice site, but rather in the effective diffusion coefficient that results from a free diffusion process that may be comprised of jumps to neighbouring lattice sites or of jumps across multiple lattice sites. We, therefore, measure long diffusive trajectories to then determine the effective diffusion coefficient  $D$  from its relation to the mean-squared displacement

$$\text{MSD}(\Delta t) = \langle (\mathbf{x}(t + \Delta t) - \mathbf{x}(t))^2 \rangle = 4D\Delta t. \quad (\text{II.40})$$

Effective energy barriers and attempt rates for the free diffusion process can then be derived via an Arrhenius plot according to Eq. II.3.

## II.4 MOLECULAR DYNAMICS

Molecular dynamics (MD) is a computational method for the simulation of the physical time evolution of interacting particles by numerically solving Newton's equations of motion [113]. While many of the numerical methods used in MD simulations have been developed much earlier, their successful application to many-body problems was only enabled by the development of modern electronic computers in the mid 20th century, after which MD became a popular simulation method [114–117] (shortly after the development of MC methods). For this thesis, we are not implementing an MD algorithm ourselves, but we are using the "Large-scale Atomic/Molecular Massively Parallel Simulator" (LAMMPS) [118] code to set up MD simulations. To do this and to be able to correctly interpret the MD simulation results, an understanding of the basic ideas behind MD simulations is necessary. The following pages are going to serve as an introduction to those ideas. As sources we are mainly using lecture notes [119], my master's thesis [86] and the text books Refs. [120, 121], which are recommended reads for further details on this topic.

### II.4.1 EQUATIONS OF MOTION

A fundamental starting point for the derivation of equations of motion for an N body system are the differential equations from the Lagrange formalism

$$\frac{d}{dt} \left( \frac{\partial \mathcal{L}(\dot{\mathbf{q}}, \mathbf{q})}{\partial \dot{\mathbf{q}}_k} \right) - \frac{\partial \mathcal{L}(\dot{\mathbf{q}}, \mathbf{q})}{\partial \mathbf{q}_k} = 0 \quad k \in \{1, 2, \dots, N\}, \quad (\text{II.41})$$

where the Lagrangian  $\mathcal{L}$  is a function of the particle coordinates  $\mathbf{q} = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N\}$  and velocities  $\dot{\mathbf{q}} = \{\dot{\mathbf{q}}_1, \dot{\mathbf{q}}_2, \dots, \dot{\mathbf{q}}_N\}$ . The dependence on these variables can be separated into

$$\mathcal{L}(\dot{\mathbf{q}}, \mathbf{q}) = \mathcal{K}(\dot{\mathbf{q}}) - \mathcal{V}(\mathbf{q}) \quad (\text{II.42})$$

with a kinetic energy term  $\mathcal{K}(\dot{\mathbf{q}})$  and a potential energy term  $\mathcal{V}(\mathbf{q})$ . For a classical N body system with particles of masses  $m_i$  and pairwise interactions  $v_{ij}(\mathbf{r}_i, \mathbf{r}_j)$  (written in cartesian coordinates  $\mathbf{r}_i$ ), the energy terms  $\mathcal{K}$  and  $\mathcal{V}$  can usually be defined as

$$\mathcal{K} = \sum_{i=1}^N \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 \quad \mathcal{V} = \sum_{i=1}^N \sum_{j=i+1}^N v_{ij}(\mathbf{r}_i, \mathbf{r}_j). \quad (\text{II.43})$$

With those definitions, the Lagrangian equations of Eq. II.41 can be written as

$$m_i \ddot{\mathbf{r}}_i = -\nabla_{\mathbf{r}_i} \mathcal{V} =: \mathbf{f}_i \quad i \in \{1, 2, \dots, N\}, \quad (\text{II.44})$$

where  $\nabla_{\mathbf{r}_i} = (\frac{\partial}{\partial r_{i,1}}, \frac{\partial}{\partial r_{i,2}}, \frac{\partial}{\partial r_{i,3}})$  is the nabla operator and the right hand sides are defined to be the interaction forces  $\mathbf{f}_i$ . The Eqs. II.44 are also known as Newton's equations of motion and in this formulation they are a set of  $3N$  second-order differential equations. Solving those equations to obtain particle trajectories  $\mathbf{r}_i(t)$  therefore requires the knowledge of  $6N$  boundary conditions, which have to be set in the form of initial coordinates and velocities. While analytic solutions to those equations are infeasible (especially for large  $N$ ), in MD simulations we can calculate the solutions numerically by applying numerical integration schemes. One such scheme will be discussed in the following after we have introduced some basic examples for the interaction potentials  $v_{ij}(\mathbf{r}_i, \mathbf{r}_j)$ .

### Interaction potentials

The derivation of interatomic or intermolecular interaction potentials is a major challenge in the field of MD simulations. Especially when parts of the system are coarse-grained, the resulting effective interactions of the coarse-grained particles can be hard to determine. However, for non-coarse-grained atomistic particles, there are a few standard interaction potentials that regularly find their uses in MD simulations. As the following potentials only depend on the center-to-center distance between the two interacting particles, we are substituting  $r := |\mathbf{r}_i - \mathbf{r}_j|$  for better readability

A very popular interaction potential to describe both the Van der Waals attraction and the close proximity repulsion (Pauli repulsion) of neutral atoms is the Lennard-Jones potential [122], commonly written as

$$v_{\text{LJ}}(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right], \quad (\text{II.45})$$

where  $\epsilon$  and  $\sigma$  are free parameters of the potential, describing the strength and the range of the interaction, respectively [Fig. 25].

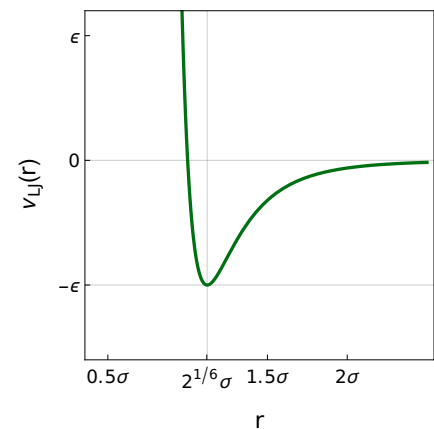


Figure 25: The Lennard-Jones interaction potential.

Whenever there are charged particles (ions, electrons or protons) involved, the Coulomb interaction [123] finds its use:

$$v_C(r) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r}, \quad (\text{II.46})$$

where  $\epsilon_0$  is the vacuum permittivity and  $q_i, q_j$  are the charges of the particles. The Coulomb interaction is repulsive if the two charges have the same sign and attractive if they have opposite signs. It is hardly used on its own, as e.g. in the attractive case a potential like this (with a singularity at distance zero) can easily lead to very strong forces and therefore to instabilities in the algorithm. For ionic systems, it can be combined with an attractive Van der Waals (for interactions with neutral particles) and a repulsive term to form a "Born-Meyer-Huggins" style potential

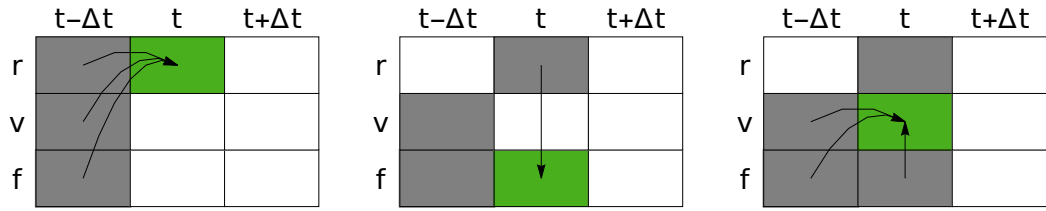
$$v_{\text{CBMH}}(r) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r} + A e^{-r/\rho} - \frac{C}{r^6}, \quad (\text{II.47})$$

where  $C$  describes the strength of the Van der Waals attraction,  $A$  and  $\rho$  the strength and range of the repulsive interaction, respectively. The potential in Eq. II.47 will for example be used in this thesis to describe the interactions of Ca and F ions in the  $\text{CaF}_2(111)$  substrate in chapter IV. The more complex coarse-grained interactions, that we are going to use for  $C_{60}$  interactions, are derived and discussed in the corresponding sections when the specific simulation setups are presented.

## II.4.2 VELOCITY VERLET INTEGRATION

When it comes to integration schemes to numerically solve the equations of motion (Eq. II.44), there is a multitude of options, each with its advantages and disadvantages. The best choice for any specific application can vary from case to case, depending on the requirements one poses to the algorithm and the resulting trajectories. The requirements on the algorithm can range from ease of implementation to the memory consumption or computational effort per time step. Concerning the resulting trajectories, their deviation from classical trajectories or their adherence to energy/momentum conservation laws, or properties like time-reversibility are things to consider. Here, we introduce a very common integrator that is a good choice for most applications and which is also used as the standard integration scheme in the LAMMPS simulation package, the velocity Verlet algorithm [124].

The velocity Verlet algorithm is used to propagate the particle coordinates  $r(t)$  and velocities  $v(t) = \dot{r}(t)$  forwards (or backwards) through time and its formulas



**Figure 26:** Illustration of the velocity Verlet algorithm's calculations for one time step.

can be derived by writing down the Taylor expansions up to the second order terms:

$$r(t + \Delta t) \approx r(t) + \dot{r}(t)\Delta t + \frac{\ddot{r}(t)}{2}\Delta t^2 \quad (\text{II.48})$$

$$v(t + \Delta t) \approx v(t) + \dot{v}(t)\Delta t + \frac{\ddot{v}(t)}{2}\Delta t^2. \quad (\text{II.49})$$

In Eqs. II.48 and II.49, we can replace  $\dot{r}(t) = v(t)$  and  $\ddot{r}(t) = \dot{v}(t) = f(t)/m$  (with  $f(t)$  being the forces that are calculated from the interaction potentials via Eq. II.44), but we need an expression to replace the  $\ddot{v}(t)$  term, which we can derive from another Taylor expansion,

$$\dot{v}(t + \Delta t) \approx \dot{v}(t) + \ddot{v}(t)\Delta t \quad \rightarrow \quad \ddot{v}(t) \approx \frac{\dot{v}(t + \Delta t) - \dot{v}(t)}{\Delta t}. \quad (\text{II.50})$$

Putting Eq. II.50 into II.49, we arrive at the formulas of the velocity Verlet algorithm:

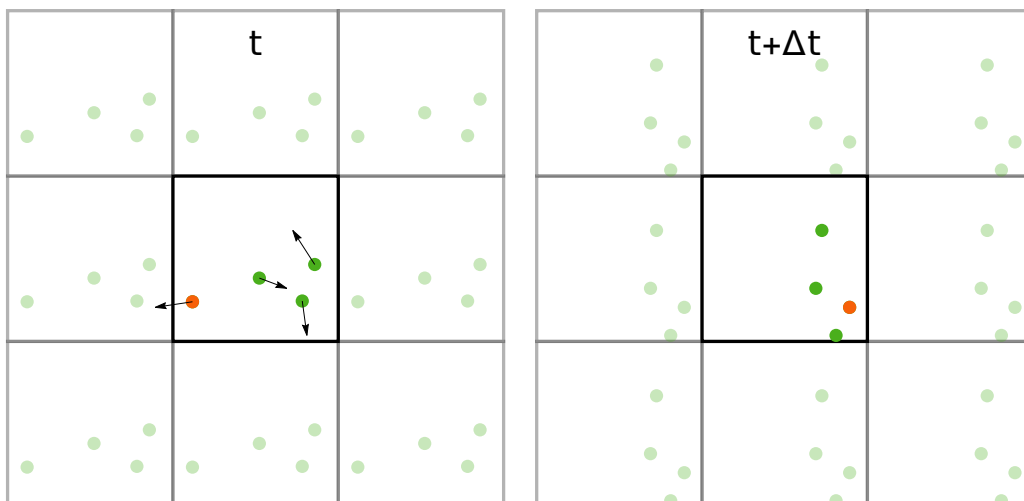
$$r(t + \Delta t) \approx r(t) + v(t)\Delta t + \frac{f(t)}{2m}\Delta t^2 \quad (\text{II.51})$$

$$v(t + \Delta t) \approx v(t) + \frac{f(t + \Delta t) + f(t)}{2m}\Delta t. \quad (\text{II.52})$$

The calculation order of this algorithm is illustrated in Fig. 26. Starting with particle coordinates  $r(t)$  and velocities  $v(t)$  at time  $t$ , we can first calculate the acting forces  $f(t)$  from  $r(t)$  using the given interaction potentials of the system. These can then be used to calculate the coordinates at the following time step  $r(t + \Delta t)$  with Eq. II.51, from which we can, in turn, calculate the forces  $f(t + \Delta t)$ . Finally, we can then calculate the velocities at the next time step using Eq. II.52.

### II.4.3 PERIODIC BOUNDARY CONDITIONS

The scope of MD and MC simulations is usually limited to a maximum of several million particles due to constraints of memory or computation time. For



**Figure 27:** Visualization of PBCs. The orange particle is highlighted, as it crosses the periodic boundary in the shown time step and ends up on the other side of the simulation box.

example, the MD simulations presented in this thesis are composed of a few hundred to a few thousand interacting particles. With so few particles, one can only construct microscopic configurations. A common trick to virtually expand the simulated configuration into an infinite bulk/plane is the use of periodic boundary conditions (PBCs). The simulation box is treated as a primitive cell that is replicated in  $x/y/z$  directions to form an infinite periodic lattice [Fig. 27]. In principle, every particle in the simulation now interacts with an infinite number of periodic images. However, the number of force calculations can be reduced to a finite number, e.g. by using a cut-off radius  $r_c$  for short-ranged interactions (like the Lennard-Jones potential) above which the interactions are truncated. For long-ranged interactions (like the Coulomb potential), shifting parts of the calculation into  $k$ -space (e.g. with so-called "Ewald sums") can reduce the computational effort [125].

#### II.4.4 INITIALIZATION AND THERMOSTATING

When performing MD simulations, we are usually interested in the properties of the system at specific temperatures. Setting up the system to be and stay at a given temperature is the first thing we have to solve. One can try to initialize the system in a state that is close to equilibrated at a given temperature by pulling random initial velocities from the Maxwell-Boltzmann distribution

$$p(\mathbf{v}) = \frac{1}{(2\pi k_B T/m)^{3/2}} e^{-\frac{m\mathbf{v}^2}{2k_B T}}, \quad (\text{II.53})$$



but the harder problem is to also put the particles into the right initial positions such that no net transfer between the time-averaged kinetic and potential energy takes place (which leads to a shift of the temperature). Instead of trying to perfectly initialize the system into an equilibrated state, it is therefore mostly easier to initialize the coordinates either on a grid or randomly and then deal with the resulting temperature shift with temperature control methods, so-called "thermostats".

### Velocity rescaling

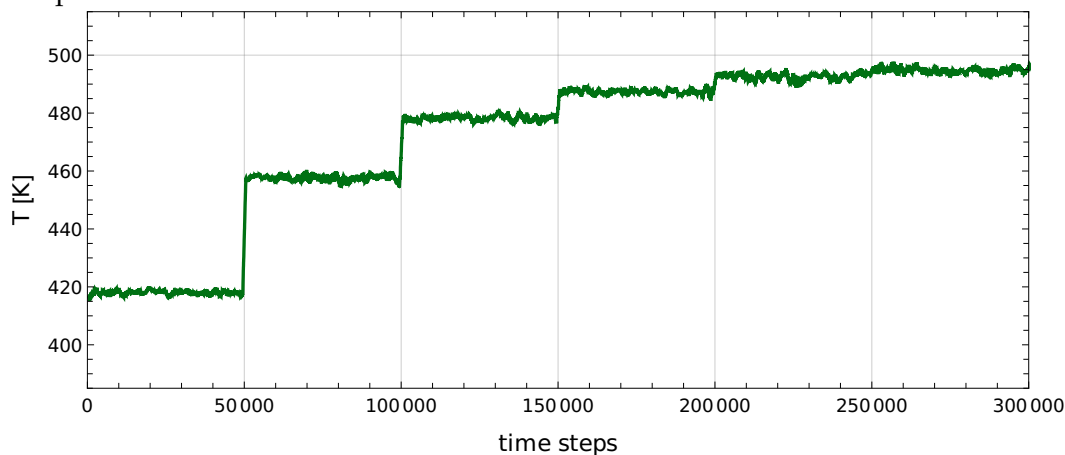
Velocity rescaling is a conceptually very simple approach to control the temperature of a system by monitoring it and then rescaling the particle velocities to produce a desired temperature. To do this, we first have to define the instantaneous temperature  $T_{\text{inst}}$  in an MD simulation via the relation between kinetic energy  $E_{\text{kin}}$  and temperature  $T$ :

$$\frac{d}{2} N k_B T = \langle E_{\text{kin}} \rangle = \sum_{i=1}^N \frac{1}{2} m_i \langle \mathbf{v}_i^2 \rangle \quad (\text{II.54})$$

$$\rightarrow T_{\text{inst}} := \sum_{i=1}^N \frac{m_i \mathbf{v}_i^2}{d N k_B}, \quad (\text{II.55})$$

where  $d$  denotes the degrees of freedom per particle ( $d = 3$  for point particles in three dimensions) and  $N$  the number of particles in the system. With this definition, the temperature  $T$  can be calculated as the time average of the instantaneous temperature,  $T = \langle T_{\text{inst}} \rangle$ . After monitoring the instantaneous

**Figure 28:** Example equilibration of a small ( $N = 288$ ) system using velocity rescaling. Starting at a temperature of 420 K, velocity rescaling is applied every 50000 time steps, resulting in a temperature close to the desired 500 K after a few applications. The displayed temperature is a running average of the instantaneous temperature.



temperature over some time to calculate a time average, we can enforce a certain desired temperature  $T_{\text{des}}$  by rescaling the particle velocities according to the transformation

$$\mathbf{v}_i \rightarrow \mathbf{v}'_i = \mathbf{v}_i \cdot \sqrt{\frac{T_{\text{des}}}{\langle T_{\text{inst}} \rangle}}. \quad (\text{II.56})$$

This will shift the temperature to  $T_{\text{des}}$ , however, such a discontinuous velocity rescaling will usually again be followed by a net transfer between potential and kinetic energy, shifting the temperature away from  $T_{\text{des}}$  over the course of the next few time steps [Fig. 28]. Therefore this velocity rescaling has to be applied multiple times, with short equilibration/temperature monitoring intervals in between. As these jumps and relaxations in temperature do not resemble the behaviour of any real physical system (the results do not correspond to a canonical NVT ensemble), velocity rescaling is not considered to be a proper thermostat. However, it is still a very useful and simple tool to equilibrate a system at the desired temperature.

### *Langevin thermostat*

The Langevin thermostat controls the systems temperature by modifying the equations of motion with a friction drag and a random force term,

$$m_i \ddot{\mathbf{r}}_i = -\nabla_{\mathbf{r}_i} \mathcal{V} - \gamma m_i \dot{\mathbf{r}}_i + \mathbf{W}_i. \quad (\text{II.57})$$

with a friction coefficient  $\gamma$  and a random force vector  $\mathbf{W}_i(t)$  with no correlations in time or between particles:

$$\langle \mathbf{W}_i(t) \cdot \mathbf{W}_j(t') \rangle = 6m_i k_B T \gamma \delta_{ij} \delta(t - t'), \quad (\text{II.58})$$

where  $\delta_{ij}$  is the Kronecker delta and  $\delta(x)$  the Dirac delta distribution. With this modification to the equations of motion, the particles will behave like they are moving in a solvent of thermostat particles. The drag term makes the particles lose kinetic energy to the solvent proportional to their velocity and the random force term periodically transfers kinetic energy from the solvent back to the particles. The resulting trajectories resemble a canonical (NVT) ensemble, however, the generated trajectories can only be representative of systems with big particles in solution (e.g. colloids). Still, the random nature of the Langevin thermostat makes it also useful in atomistic and molecular simulations for the generation of random initial configurations at any given temperature, e.g. from an initially regular grid of particles.

## II.4.5 GENERIC SIMULATION PROCEDURE

Putting the discussed MD simulation parts together, we can outline a generic simulation procedure that is applicable to the observation of surface diffusion processes:

1. Set up an initial configuration in which the particles are placed on a regular grid that closely represents the crystal structure of the materials of interest.
2. Equilibration:
  - Start with a short run in which the Langevin thermostat with a unique random seed is applied to give the particles some random initial velocities.
  - Turn off the Langevin thermostat and start the equilibration run in which velocity rescaling is applied multiple times until the temperature of the system is sufficiently close to the desired temperature.
  - Turn off the velocity rescaling
3. Measurement:
  - Let the system evolve without the application of any temperature control.
  - After predefined time intervals or after a transition was detected, save the current configuration into an output file for later analysis.
  - (optional) Monitor and output observables like potential energy / temperature in predefined time intervals.

As already mentioned, our MD simulations are all implemented using the LAMMPS simulation package, which comes with an outstanding amount of functionality out of the box. All the common thermostats and interaction potentials are implemented and custom potentials can be supplied to LAMMPS in tabulated form. Instructions can be given to the program in the form of an input script. An example input script is given in the appendix [B.4](#)



# III

## $C_{60}$ ON $C_{60}(111)$

### CONTENTS

---

III.1 MD Simulations . . . . .	45
III.1.1 The Girifalco Potential . . . . .	45
III.1.2 Simulation Setup . . . . .	48
III.1.3 Raw Transition Rate Measurements . . . . .	49
III.1.4 Transition Rate Modelling and Discussion . . . . .	52
III.2 KMC Simulations . . . . .	56
III.2.1 Tuning of the "Simple" model . . . . .	56
III.2.2 Multilayer Growth . . . . .	61
III.3 Summary . . . . .	66

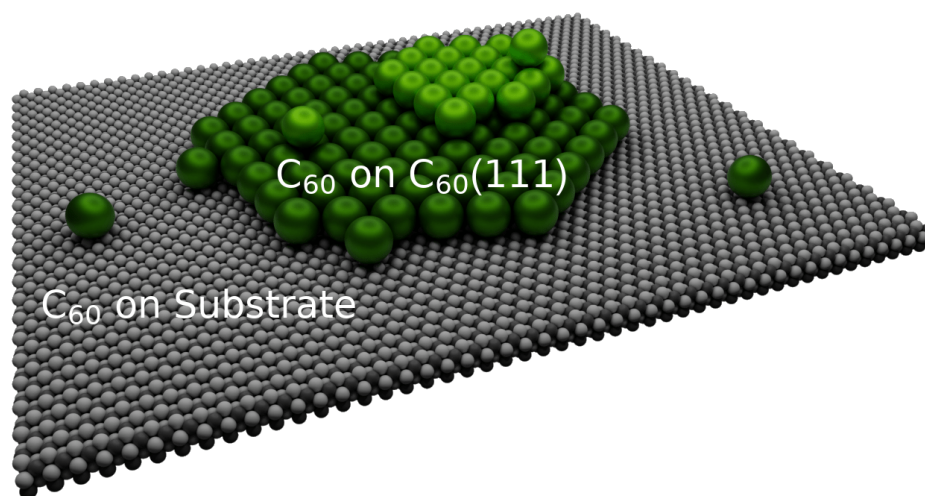
---

In this chapter, we investigate the epitaxial growth process of  $C_{60}$  on an ideal  $C_{60}(111)$  crystal surface. While there are no experimental observations on this exact system, a well-established understanding of it is transferable to other systems like the growth of multilayered  $C_{60}$  clusters on substrates like  $CaF_2(111)$ , where the first monolayer of  $C_{60}$  molecules grows closely packed to form a  $C_{60}(111)$  surface [Fig. 29]. A model for the transition rates of  $C_{60}$  on  $C_{60}(111)$  diffusion processes is therefore valid for the processes on the second layer and onwards even in the  $C_{60}$  on  $CaF_2(111)$  system. The contents of this chapter are largely based on our first publication (Ref. [1]).

In the first part of this chapter, we detail how we measure transition rates of  $C_{60}$  on a clean  $C_{60}(111)$  crystal surface in a wide variety of initial configurations and temperatures in MD simulations. From those transition rate measurements, we derive energy barriers and attempt rates from Arrhenius plots of the individual transition types. On that basis, we construct a "RawMD" rate model that simply aims to reproduce the MD measurements.

In the second part of this chapter, we then show that this naive approach leads to a thermodynamically inconsistent rate model, manifesting in a net entropy production in KMC simulations even when the system is in a steady state. We then reduce the model to a "Simple" version that is thermodynamically consistent and generates the same cluster morphologies. Finally, this Simple model is then used to run  $C_{60}$  on  $C_{60}(111)$  multilayer growth KMC simulations, comparing our results with the work of Bommel et al [55]. We are going to uncover, that the high cluster densities that have been observed in this particular study of  $C_{60}$  multilayer growth on mica can be explained by the formation of grain boundaries on the coalescence of the first monolayer.

**Figure 29:** On many substrates (like on  $CaF_2(111)$ ),  $C_{60}$  layers grow closely packed, forming a  $C_{60}(111)$  surface.



## III.1 MD SIMULATIONS

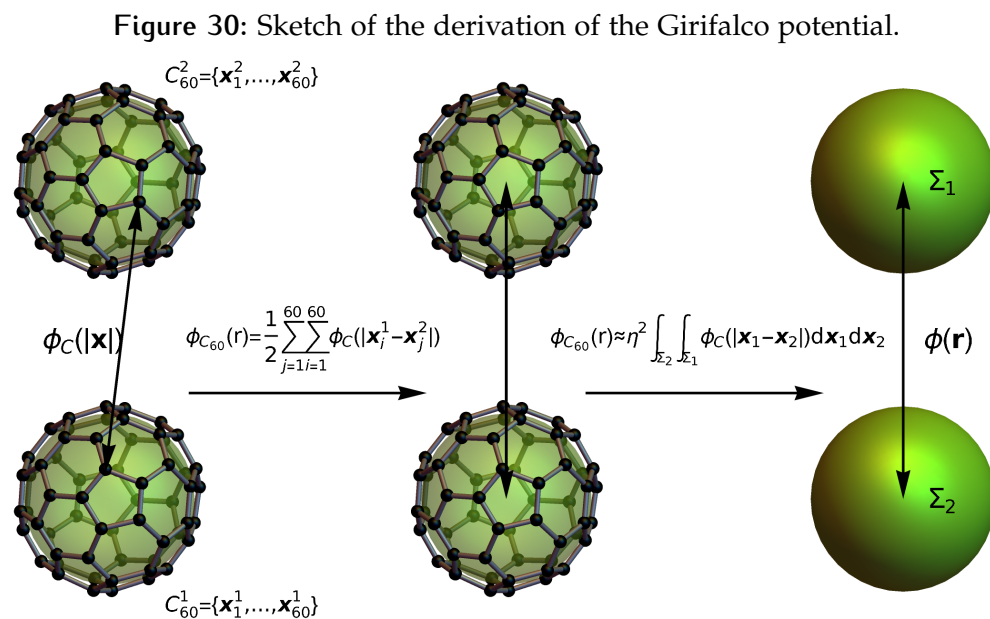
### III.1.1 THE GIRIFALCO POTENTIAL

We start with the introduction of the Girifalco potential, a coarse-grained potential that simplifies the  $C_{60}$ - $C_{60}$  interaction by reducing each  $C_{60}$  molecule to a single (spherically symmetric) bead. It was first derived in 1991 by L. A. Girifalco [22, 23]. In a pure  $C_{60}$  system, using such a coarse-grained potential reduces the number of force calculations per pair of molecules by a factor of  $60^2 = 3600$  and also allows for the use of a larger time step in the MD simulation, enabling us to observe processes that take place on very long time scales like edge diffusion processes with up to four initial neighbours.

To derive the potential, we start by assuming a Lennard-Jones interaction  $\phi_C(r)$  between the individual carbon atoms of different  $C_{60}$  molecules, which we define as

$$\phi_C(r) = -\frac{A}{r^6} + \frac{B}{r^{12}}. \quad (\text{III.1})$$

In an atomistic representation of the  $C_{60}$  molecule, the overall interaction between two  $C_{60}$  molecules is then given by the sum over all pairwise interactions of the individual C atoms of the two different molecules. To get to the Girifalco



potential, we approximate this sum by integrating over the surfaces of two spheres  $\Sigma_1$  and  $\Sigma_2$ ,

$$\phi_{C_{60}}(r) \approx \eta^2 \int_{\Sigma_1} \int_{\Sigma_2} \phi_C(r) d\Sigma_2 d\Sigma_1, \quad (\text{III.2})$$

where  $r$  is the distance between the centers of the two spheres and  $\eta = 60/4\pi R^2$  is the surface density of C atoms ( $R \approx 0.355$  nm is the  $C_{60}$  center-to-nucleus radius). Evaluating this integral, we obtain the Girifalco potential,

$$\begin{aligned} \phi_{C_{60}}(s) \approx & -\alpha \left( \frac{1}{s(s-1)^3} + \frac{1}{s(s+1)^3} - \frac{2}{s^4} \right) \\ & + \beta \left( \frac{1}{s(s-1)^9} + \frac{1}{s(s+1)^9} - \frac{2}{s^{10}} \right), \end{aligned} \quad (\text{III.3})$$

with the variable  $s = r/(2R)$  and constants  $\alpha = 60^2 A / (12(2R)^6)$  and  $\beta = 60^2 B / (90(2R)^{12})$ . Knowing the nearest neighbour distance  $d_0 \approx 1.0$  nm and the sublimation energy  $E_{\text{Sub}} \approx 1.8$  eV of a  $C_{60}$  fcc crystal, we can derive values for the constants  $\alpha$  and  $\beta$  from the conditions

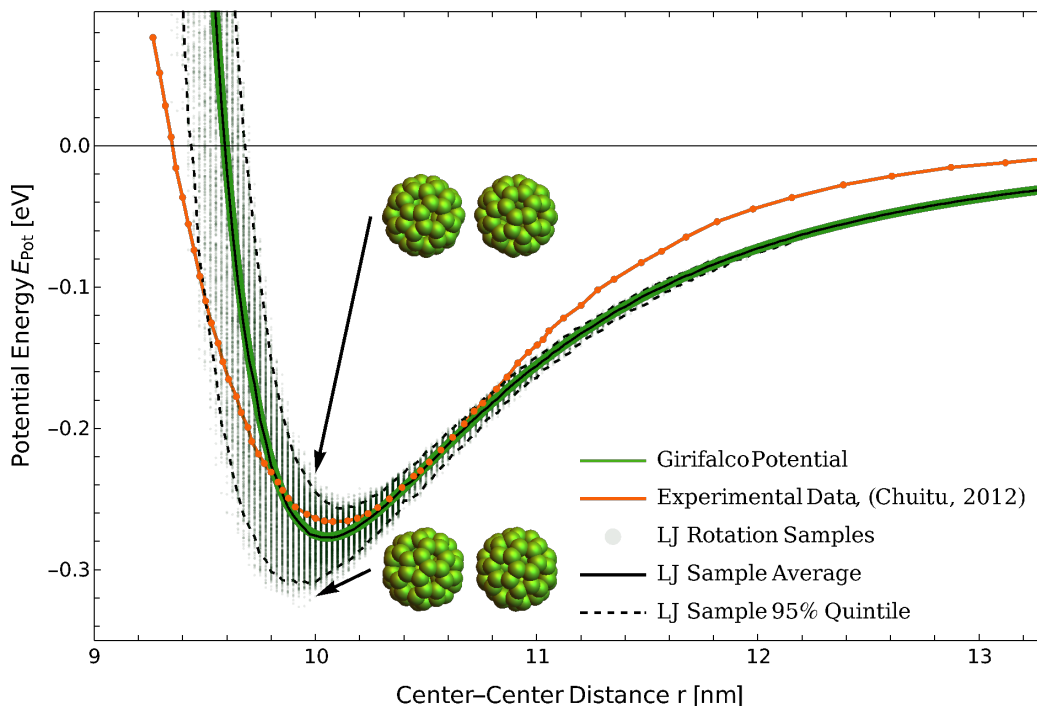
$$E := \sum_{r \in \Lambda_{\text{fcc}}} \phi_{C_{60}}(s) \stackrel{!}{=} 2E_{\text{Sub}} \quad \frac{\partial \phi_{C_{60}}}{\partial r}(d_0) \stackrel{!}{=} 0. \quad (\text{III.4})$$

The first condition of Eq. III.4 sums over the positions of an fcc crystal  $\Lambda_{\text{fcc}}$  that is constructed around a center probe with nearest neighbour distance  $d_0$ . We obtain the parameters

$$\alpha = 46.8 \times 10^{-3} \text{ eV} \quad \beta = 84.6 \times 10^{-6} \text{ eV}. \quad (\text{III.5})$$

A comparison of this Girifalco potential with an experimentally measured  $C_{60}$ - $C_{60}$  interaction potential [76] is shown in Fig. 31. The agreement of the experimental potential with the Girifalco potential is relatively good around the energy minimum but deviates pretty strongly for the short-range repulsion and the long-range attraction. However, considering that this is just one measurement with one specific orientation between the two molecules and also that there might be other systematic errors in the measurement, the agreement is remarkable. Also shown in Fig. 31 are samples of the potential energy between two atomistic  $C_{60}$  representations with random orientations towards one another, interacting via the Lennard-Jones potential of Eq. III.1. It can be seen how the sample average aligns perfectly with the Girifalco potential, the expected effect of integrating out the rotational degrees of freedom. The samples and their 95% quintiles also hint at the limitations of the Girifalco potential as they show how atomistically represented  $C_{60}$  molecules can actually form stronger bonds (at a shorter center-to-center distance) by aligning





**Figure 31:** Comparison of the Girifalco potential with experimental data [76] and with samples of randomly oriented atomistic  $C_{60}$  molecules interacting via the Lennard-Jones potential III.1. The insets show the maximum and minimum energy configurations of the samples at  $r = 10$  nm.

themselves correctly. Because of this, the Girifalco potential is usually only assumed to be a good approximation for temperatures well above 260 K. At 260 K, crystalline  $C_{60}$  goes through a structural phase transition from a phase in which molecular orientations align (at lower temperatures) to a plastic crystal in which orientations are disordered (at higher temperatures)[16–21]. The Girifalco potential is therefore mostly used for high-temperature studies – like the prediction of a stable liquid  $C_{60}$  phase [126–130] or the analysis of  $C_{60}/C_{70}$  mixtures [131, 132]. However, the simplicity of the potential occasionally motivates its use at lower temperatures [52, 133–136]. For  $C_{60}$  in bulk, a comparative study between an atomistic and coarse-grained (Girifalco) representation of the molecule was done in Ref. [137] employing MD simulations in a temperature range of  $T \in (300, 1900)$ K and a range of different pressures, showing that the Girifalco potential yields similar results as the atomistic representation for low (ambient) pressures.

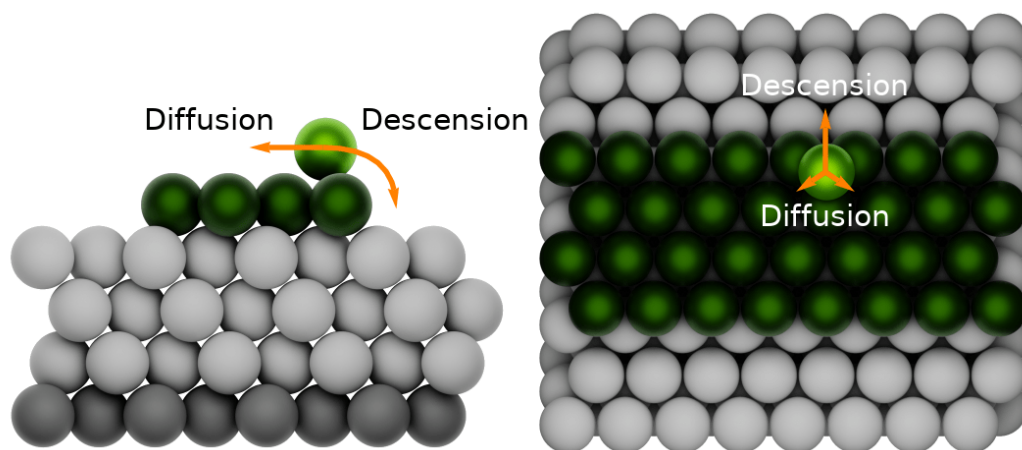
In the simulations of this chapter, we are neither using high pressures nor low temperatures, so the Girifalco potential is a reasonable choice. However, when using the resulting rate model in KMC simulations, one has to keep in mind that the rates may be less accurate in a low-temperature regime.

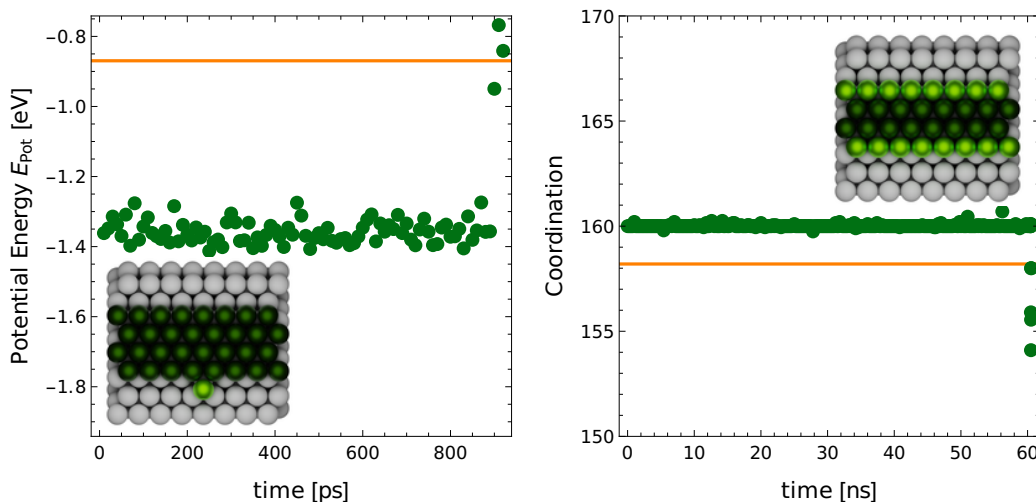
### III.1.2 SIMULATION SETUP

Using LAMMPS and a tabulated version of the Girifalco potential, we set up MD simulations of interacting  $C_{60}$  molecules. The  $C_{60}(111)$  substrate is modelled as a four layer thick crystal of which the bottom layer is fixed in place. On top of this crystal, an environment of deposited particles in a stable configuration is set up for one or a few tagged particles that are positioned in metastable states in which transitions of interest can occur. We measure the lifetime of those metastable states as well as the frequency of outgoing transition types  $i$  to determine the associated transition rates  $k_i$ . An example setup of such an MD simulation is shown in Fig. 32. To equilibrate the system to a random starting configuration at a given temperature, a combination of the Langevin thermostat and velocity rescaling is applied before the measurement begins. After equilibration, the lifetime of the initial state is measured in an NVE simulation (thermostat and velocity rescaling turned off).

We run a large number of trajectories for every configuration to gather enough statistics on all possible transitions. The temperature range varies depending on the configuration but overall the used temperature range is  $T \in (200, 850)$ K. The initial configurations can be categorized by the initial number of neighbours  $n$  of their tagged particles, which we vary from  $n = 0$  (free diffusing particle) to  $n = 4$ . To notice if a transition has occurred, the potential energy of the tagged particles or their coordination number with particles in their neighbourhood are

**Figure 32:** Example MD simulation configuration for the measurement of the descension rate. The bottom layer of  $C_{60}$  molecules (gray) is fixed in place to ensure an fcc crystal structure with a (111) surface. The three layers of white-coloured  $C_{60}$  molecules evolve freely during the NVE simulation but are under the effect of a Langevin thermostat or velocity rescaling during equilibration. The dark green particles are put into a stable configuration where they are unlikely to leave their position and set up the environment for the tagged light green particle, which is put into a state where the transitions of interest can occur.





**Figure 33:** Potential energy and coordination as stop triggers to notice transitions. Left: Potential energy over time of a single particle bound to a slab of particles in a two-neighbour configuration, with a threshold (orange) set to notice the dissociation from - or ascension onto - the slab. Right: Coordination number of the particles at the edge of a slab (in four-neighbour configurations) with a threshold (orange) set to notice when one of the particles pops out of - or ascends onto - the slab

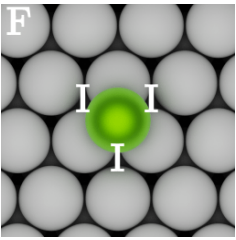
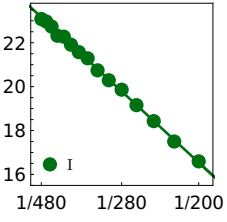
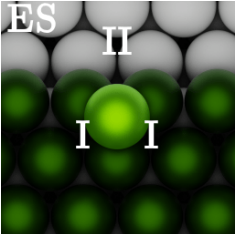
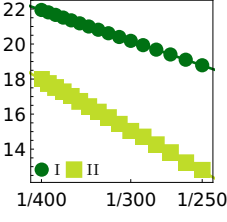
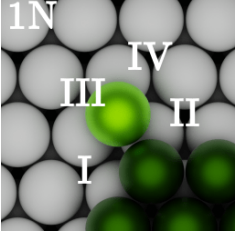
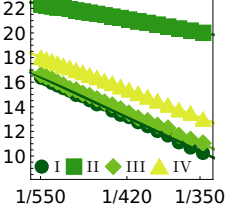
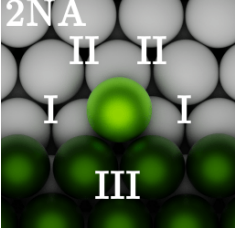
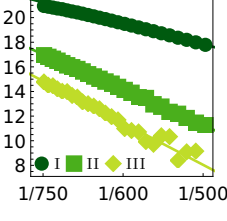
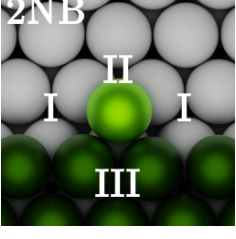
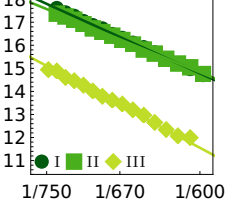
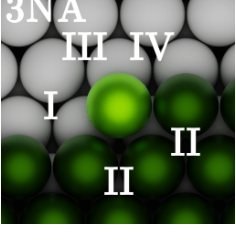
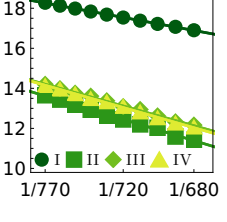
measured throughout and used as triggers to stop the simulation when they cross some threshold value. Examples are shown in Fig. 33 for particles in two-neighbour and four-neighbour initial configurations.

### III.1.3 RAW TRANSITION RATE MEASUREMENTS

First, we are going to present the transition rate measurements with as little modelling as possible. In Tab. 1 and 2 the implemented initial configurations are shown together with possible transition targets and configuration labels. Suppose we have gathered  $N(T)$  trajectories of a certain initial configuration at temperature  $T$  and have observed  $N_i(T)$  occurrences of transition target  $i$ , we estimate the transition rate for transitions to target  $i$  via

$$k_i(T) \approx \frac{N_i(T)}{t_{\text{tot}}(T)} \quad t_{\text{tot}}(T) = \sum_{j=1}^{N(T)} \tau_j, \quad (\text{III.6})$$

where  $\tau_j$  are the run times of each of the trajectories until a transition occurred. In the case of free diffusion (F), the transition rates were derived from the mean-squared displacement (MSD) of the diffusion trajectories as described in the methods section II.3.5. In the Arrhenius plots shown in Tab. 1 and 2 we can nicely see how all the calculated transition rates obey the Arrhenius law by

Configuration	Arrhenius Plot	T	Target i	$\Delta E_i$ [meV]	$\nu_{0,i}$ [THz]
		200 — 480 K	I	195.6(20)	0.43(4)
		250 — 400 K	I II	181.4(3) 300.2(13)	0.325(3) 0.385(18)
		350 — 550 K	I II III IV	501(4) 188.3(6) 409(2) 458(4)	0.53(6) 0.262(4) 0.38(2) 0.25(4)
		500 — 750 K	I II III	439(2) 744(5) 879(34)	0.74(3) 1.23(12) 2.29(144)
		600 — 750 K	I II III	779(5) 708(4) 917(29)	4.60(38) 2.46(18) 5.33(266)
		680 — 770 K	I II III IV	686(5) 1119(30) 1037(20) 1062(24)	1.33(10) 4.32(210) 4.55(148) 6.29(165)

**Table 1:** Summary of the results from our MD Simulations of the  $C_{60}$  on  $C_{60}(111)$  system. The first column shows the configuration with tagged particle (light green) and the possible transitions (white numerals). The second column shows the corresponding Arrhenius plot obtained from the MD data ( $1/T$  with  $T$  in Kelvin on the x-axis,  $\log(k)$  with  $k$  in Hertz on the y-axis).

Configuration	Arrhenius Plot	T	Jump to	$\Delta E_i$ [meV]	$\nu_{0,i}$ [THz]
		680	I	989(17)	3.01(82)
		—	II	1119(24)	4.76(186)
		770	III	967(13)	3.49(73)
		K	IV	618(9)	0.20(3)
		750	I	906(12)	0.91(16)
		—	II	1295(46)	2.47(169)
		840			
		K			
		750	I	1270(18)	10.4(28)
		—	II	1311(39)	3.83(217)
		840			
		K			

Table 2: Continuation of Tab. 1.

falling onto lines, so we can derive energy barriers  $\Delta E_i$  and attempt rates  $\nu_{0,i}$  (as explained in Sec. II.3.5), which are also listed in these summarizing tables. For the following discussions we are going to refer to specific transitions of Tab. 1 and 2 via a combination of the initial configuration label with the transition target label (e.g. (F)I for the free diffusion transition).

To validate our methodology we can compare some of our results to previous estimates from similar computational methods. Our value for the diffusion barrier (from config (F)I) of  $\Delta E_{FD} = 195.6(20)$  meV falls right into a range of other values in the literature: Gravil *et al.* [138] (168 meV from pair potential calculations), Liu *et al.* [52] (178(4) meV from MD simulations like ours), Cantrell and Clancy [54] (205(22) meV from molecular mechanics calculations) and Goose *et al.* [139] (207 meV from DFT calculations). The corresponding attempt rate  $\nu_{0,FD} = 0.43(4)$  THz is also close to the estimate from MD simulations by Liu *et al.* [52] of 0.20(3) THz. Another characteristic energy that we can extract from our measurements is the Ehrlich-Schwöbel barrier  $E_{ES}$ , by subtracting the free diffusion barrier  $\Delta E_{FD}$  from the descension barrier  $\Delta E_{Des}$  of transition (ES)II,

$$E_{ES} = \Delta E_{Des} - \Delta E_{FD} = (300.2 - 195.6)\text{meV} = 104.6(24)\text{meV}. \quad (\text{III.7})$$

This value is also in very good agreement with the estimate by Goose [139] of 104 meV and with the value obtained by Cantrell and Clancy of 129(30) meV. To conclude our comparison with literature values, the transitions for which comparable energy barrier values exist in the works of Cantrell/Clancy and Liu are listed in Tab. 3, all showing relatively good agreement with our results.

Transition	this work	MM [54]	MD [52]
(F)I	195.6(20)	205(22)	178(4)
(1N)III	409(2)	448(25)	429(57)
(2NB)II	708(4)	717(29)	-
(ES)II	300.2(13)	334(20)	-

**Table 3:** Comparison of energy barriers  $\Delta E_i$  (in meV) with previous results from molecular mechanics (MM) and molecular dynamics (MD). Note that the MM results are consistently larger.

### III.1.4 TRANSITION RATE MODELLING AND DISCUSSION

For the modelling of the measured transition rates we are going to start with the assumption that the energy barriers and attempt rates only depend on the transition type ( $i \in \{\text{FD, Diss, ED-A, ED-B, Asc, Desc}\}$ ) and the number of initial (lateral) neighbours  $n$  as motivated in Sec. II.3.3,

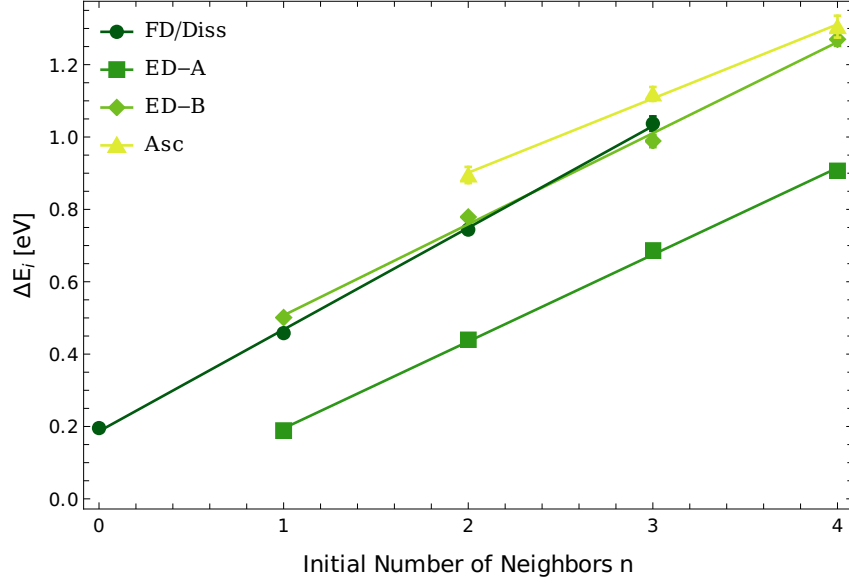
$$k_i(n, T) = \nu_{0,i}(n) e^{-\Delta E_i(n)/k_B T}. \quad (\text{III.8})$$

The energy barriers that we have extracted from the MD results for the various transition types are plotted in Fig. 34 as functions of the initial number of neighbours. In this plot, we can directly see a linear dependence of the energy barriers to  $n$ , which motivates a parametrization of the energy barriers with the bond counting approach,

$$\Delta E_i(n) = \Delta E_{i,0} + (n - \hat{n}_i) E_{B,i}, \quad (\text{III.9})$$

with a base energy barrier  $\Delta E_{i,0}$ , number of sustained bonds  $\hat{n}_i$  in transitions of type  $i$ , and an effective bond energy  $E_{B,i}$  that can also depend on the transition type  $i$ .

The linear fits that are shown in Fig. 34 give us values for the base energy barriers  $\Delta E_{i,0}$  and effective bond energies  $E_{B,i}$ , which are listed in Tab. 4. The energy minimum of the used Girifalco potential is  $E_G = 277$  meV and is



**Figure 34:** Energy barriers plotted against the number of initial neighbours for the various observed transition types. The FD/Diss barriers are taken from transitions (F)I, (1N)IV, (2NA)II and (3NA)III (See Tab. 1 and 2), combining the free diffusion barrier with the long dissociation paths that represent a complete dissociation from the cluster. The ascension barriers do not show a significant difference between A and B step edges, so the shown data points are calculated as averages of the measurements at the two respective edges. The data points for ED-A are taken from transitions (1N)II, (2NA)I, (3NA)I and (4NA)I, while the data points for ED-B are given by transitions (1N)I, (2NB)I, (3NB)I and (4NB)I. All data sets are well described by the plotted linear fits.

the expected energy barrier for when a bond has to be broken completely. This expectation is fulfilled for the case of free diffusion/dissociation, where the fitted effective bond energy of  $E_{B,FD/Diss} = 276(5)$  meV is very close to  $E_G$ . For the case of edge diffusion at A and B step edges, we observe that the effective bond energies overlap within their margins of error and are both significantly lower than the full bond energy  $E_G$ . This is an expected limitation of the bond counting approach as the transition states of the edge diffusion transitions are

Transition type	$i$	$E_{B,i}$ [meV]	$\Delta E_{i,0}$ [meV]
free diffusion/dissociation	FD/Diss	276(5)	192(6)
A step edge diffusion	ED-A	245(4)	190(5)
B step edge diffusion	ED-B	254(10)	508(13)
ascension	Asc	206(12)	901(14)
descension	Desc	276(5)	300(1)

**Table 4:** Fitted effective bond energies  $E_{B,i}$  and base energy barriers  $\Delta E_{i,0}$ .

close enough to the initial neighbours such that the transitioning particle does not have to overcome the full bond energy to slide into the target state. This serves as a reminder that, while the  $C_{60}$ - $C_{60}$  interaction range is relatively short, it is still long enough that a bond counting approach that only considers the initial number of neighbours has to be viewed as a rough approximation. A very surprising result is the significantly lower effective bond strength  $E_{B,Asc} = 206(12)$  meV of the ascension transition, which is so far not fully understood. An analysis of the exact transition paths of ascension may lead to interesting insights but is not performed here as the ascension barrier of  $C_{60}$  on  $C_{60}(111)$  is high enough for those transitions to not be very relevant for the following KMC simulations. For the case of descension we only measured one initial configuration ( $n = 0$ ), so we assume that  $E_{B,Desc}$  is equal to  $E_{B,FD/Diss}$  as the descension transitions also have long transition paths in which all initial bonds have to be overcome completely.

Turning to the base energy barriers  $\Delta E_{i,0}$ , it is notable that the base energy barrier for ED-A (190(5) meV) turns out to be very close to the one for FD/Diss (192(6) meV) because the transition path for ED-A is basically initialized by a small diffusion hop to a metastable state (e.g. see transition (1N)II of Tab. 1). The base energy barrier for ED-B (508(13) meV) is more than a bond energy larger than the one of ED-A, a significant difference that is important for the evolution of the triangular star-shaped clusters (see Sec. II.2.2) that have also been observed in experiments [33–36, 42, 45]. The base energy barrier for ascension (901 meV) is very close to the adsorption energy of  $C_{60}$  on  $C_{60}(111)$  (around 930 meV, calculated with the Girifalco potential at zero Kelvin), in line with the expectation that the full adsorption energy has to be overcome to ascend to a higher layer.

The measured attempt rates do not exhibit any specific behaviour depending on the number of initial neighbours so we do not add any model on top of them. The exact values of the attempt rates  $\nu_{0,i}(n)$  are summarized in Tab. 5.

Transition type	i	Initial neighbours n				
		0	1	2	3	4
free diffusion/dissociation	FD/Diss	0.43	0.25	1.23	4.02	-
A step edge diffusion	ED-A	-	0.26	0.74	1.33	0.91
B step edge diffusion	ED-B	-	0.53	4.60	3.01	10.4
ascension	Asc	-	-	3.81	4.54	3.15
descension	Desc	0.38	0.38	0.38	0.38	0.38

Table 5: Effective attempt rates  $\nu_{0,i}(n)$  (in units of THz).



The magnitude of those attempt rates ( $10^{11} - 10^{13}$  Hz) lies in the same range as other previously measured values for large molecules from the literature [24, 140, 141].

In the KMC simulations of the following sections, we are going to consider two rate model variants. The first one is termed the "RawMD" model and is entirely based on the energy barriers and attempt rates that we have just discussed and summarized in Tab. 4 and 5. The second one – called the "Simple" model – will restrict the free parameters of the rate model a lot, reducing the attempt rate parameters and the effective bond strengths to just a single value,  $\nu_{0,i}(n) = \nu_0$  and  $E_{B,i} = E_B$ . In total the RawMD model ends up having effectively 20 attempt rate and 10 energy barrier parameters, while the Simple model only has one attempt rate and six energy barrier parameters.

## III.2 KMC SIMULATIONS

### III.2.1 TUNING OF THE "SIMPLE" MODEL

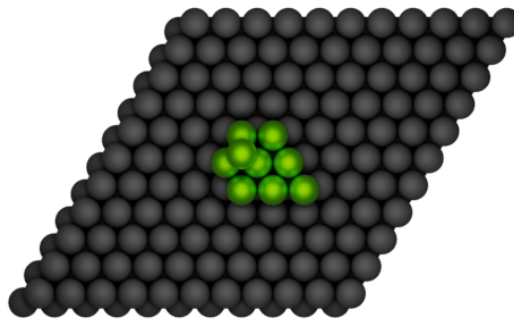
In this section, we are first going to show that the naive approach of building a rate model only focused on reproducing MD simulation results (RawMD) does not produce a thermodynamically consistent model, while on the other hand, the Simple rate model will prove to be thermodynamically consistent. We are then going to compare the cluster morphologies that the two models produce in KMC simulations and find fixed values for  $\nu_0$  and  $E_B$  with which the Simple model comes close to the RawMD results.

#### *Detailed Balance and Entropy Production*

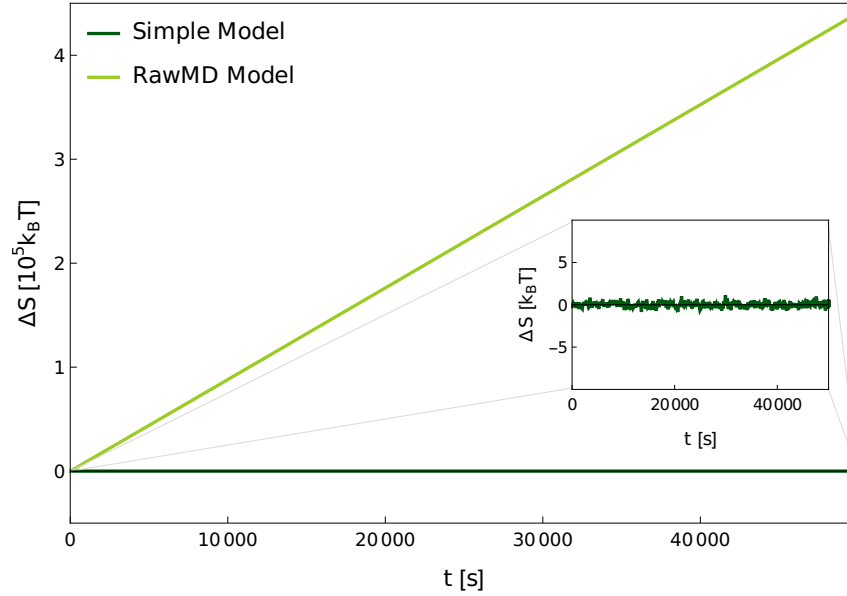
To evaluate the thermodynamic consistency of our models, we are going to employ a detailed balance condition. While the atomistic dynamics of the MD simulations are derived from a Hamiltonian, therefore obeying detailed balance and guaranteeing an absence of steady dissipation, the rate model we construct based on those dynamics does not necessarily conserve those properties. In KMC simulations, the detailed balance condition can be expressed by  $p_S k_{S \rightarrow S'} = p_{S'} k_{S' \rightarrow S}$ , where  $S$  and  $S'$  are states of the system (given by positions of all the molecules),  $p_S$  and  $p_{S'}$  are stationary probabilities to observe those configurations and  $k_{S \rightarrow S'}$  /  $k_{S' \rightarrow S}$  are the transition rates leading from one configuration to the other. While we do not have direct access to the stationary probabilities, we do have the transition rates  $k_i(n, T)$  (Eq. III.8) of which we want to know if they obey detailed balance with the extracted energy barriers and attempt rates.

A practical way to test this is to calculate the stochastic entropy production [142, 143] of the KMC trajectories. For that purpose, we prepare a  $10 \times 10$  unit cell system (cf. Fig. 35) containing nine particles assembled into a cluster (without deposition of additional particles) and run trajectories at a temperature of  $T = 318$  K. In the resulting trajectories, the particles go through the various transition types in different successions and we

can observe if any net entropy production occurs by calculating the change of entropy after every executed KMC step as the logarithm of the ratio between the rate  $k_i(n, T)$  of the transition that occurred and the rate  $k_j(m, T)$  ( $m$  being



**Figure 35:** Entropy production measurement system.



**Figure 36:** Comparison of the average entropy production of the RawMD and the Simple model over time. The average is calculated from an ensemble of 100 trajectories at temperature  $T = 318$  K.

the new number of neighbour) of the transition that would reverse the KMC step,

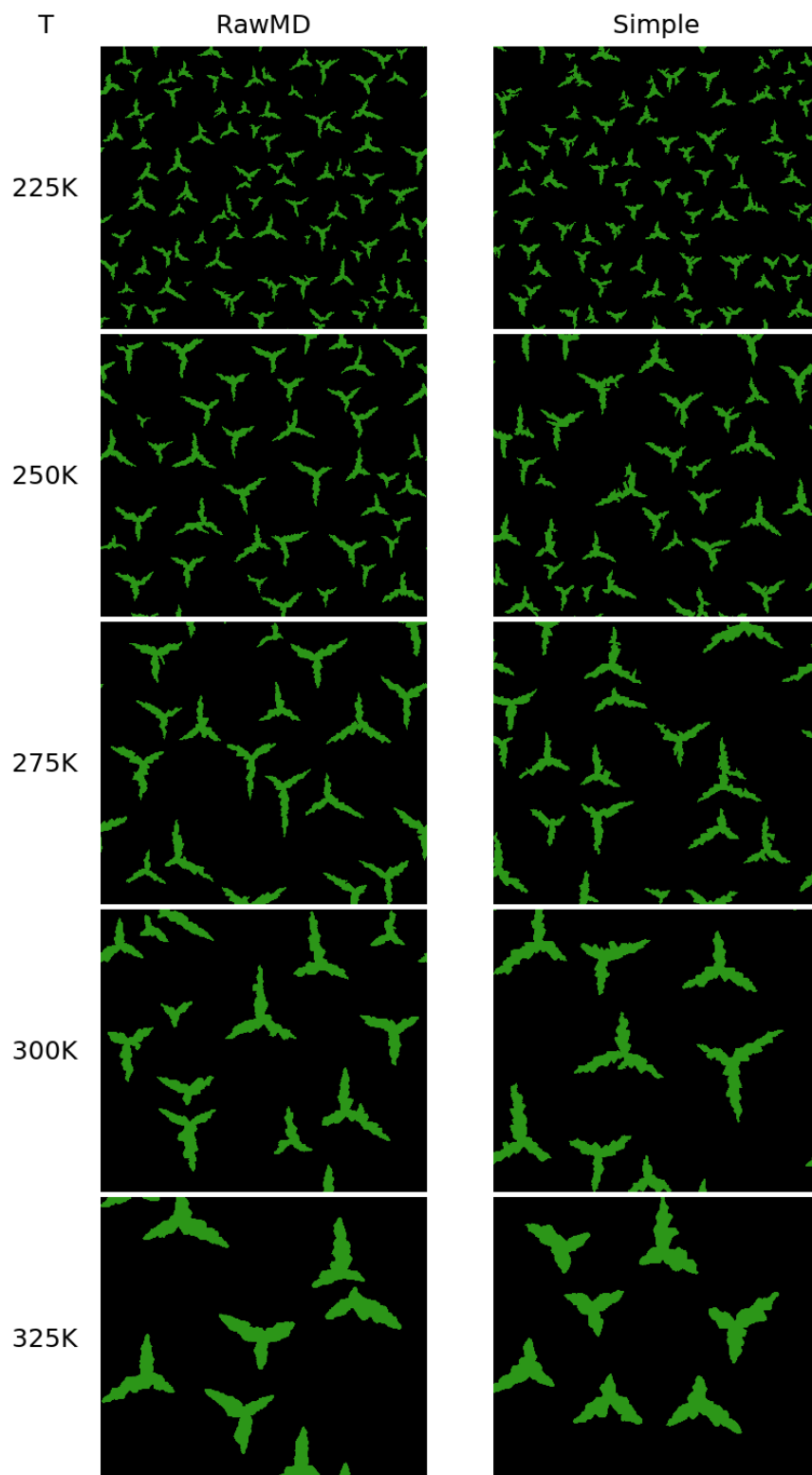
$$\delta s = \log \left( \frac{k_i(\mathbf{n}, T)}{k_j(\mathbf{m}, T)} \right). \quad (\text{III.10})$$

The sum of these single contributions  $\delta s$  over the course of the trajectories is the total entropy production  $\Delta S$ . After an initial relaxation phase, this entropy production  $\Delta S$  should fluctuate around zero if detailed balance is obeyed. Plotted in Fig. 36 is an ensemble average of  $\Delta S$  over 100 trajectories for both of our models, showing that the RawMD model exhibits a substantial dissipation of  $8.81k_B T$  of entropy per second. In contrast, the entropy production in the Simple model perfectly fluctuates around zero, showing that it obeys detailed balance.

This analysis demonstrates how a naive transfer of transition rates from MD simulations can easily yield a thermodynamically inconsistent model and that it is necessary to impose some modelling constraints (as discussed in Sec. II.3.3) onto the derived rate model to ensure thermodynamic consistency.

### *Cluster Morphologies*

To find suitable parameters  $\nu_0$  and  $E_B$  for our Simple model we are measuring and matching the cluster morphologies that the two models are producing in a temperature range of  $T \in [225, 360]$  K. We prepare an empty  $1000 \times 1000$  unit cell systems (totalling an area of  $A \approx 0.866 \mu\text{m}$ ) and have a particle flux of



**Figure 37:** Snapshots from KMC simulations using the RawMD and Simple model at several temperatures. The shown areas span  $1000 \text{ nm} \times 866 \text{ nm}$  and contain 100000 molecules, corresponding to 10% coverage. The parameters used in this Simple model are  $E_B = 235 \text{ meV}$  and  $\nu_0 = 0.25 \text{ THz}$ .

$F = 0.026 \text{ ML/min}$  active until a coverage of 10% is reached (taking a deposition time of roughly  $t_{\text{Depos}} \approx 230 \text{ s}$ ). The KMC simulation then continues without deposition over a relaxation time of  $t_{\text{Relax}} = 2t_{\text{Depos}} \approx 460 \text{ s}$  during which the clusters can relax their particle positions.

Snapshots of five different temperatures taken at the end of the simulations are shown in Fig. 37 for the RawMD model and a Simple model with parameters  $E_B = 235 \text{ meV}$  and  $\nu_0 = 0.25 \text{ THz}$ , which turned out to produce very similar results. No significant difference between the two models can be spotted with the bare eye.

For a more quantitative test, we are calculating three geometric properties for the observed cluster morphologies: The covered area  $A$ , the cluster perimeter  $P$  and the mean border curvature  $\bar{\kappa}$  [Fig. 38]. The Area and perimeter of a cluster can be calculated from the number of molecules in the cluster  $N$  and the number of molecules at the cluster's edge  $N_E$  via

$$A = N \cdot 0.866 \text{ nm}^2 \quad P = N_E \cdot 1 \text{ nm}. \quad (\text{III.11})$$

To calculate the mean curvature  $\bar{\kappa}$  we make use of the formula for the curvature of a continuous plane curve in two dimensions given in parametric form  $\gamma(t) = (x(t), y(t))$ ,

$$\kappa(t) = \frac{|\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t)|}{(\dot{x}(t)^2 + \dot{y}(t)^2)^{\frac{3}{2}}}. \quad (\text{III.12})$$

The cluster perimeters are not given as a continuous curve, but as a set of  $N_E$  discrete border molecule positions  $x_n, y_n$ , so we need to approximate the derivatives at molecule  $n$  via finite differences

$$\dot{x}_n = \frac{x_{n+1} - x_{n-1}}{2} \quad \ddot{x}_n = \frac{x_{n+2} - 2x_n + x_{n-2}}{4} \quad (\text{III.13})$$

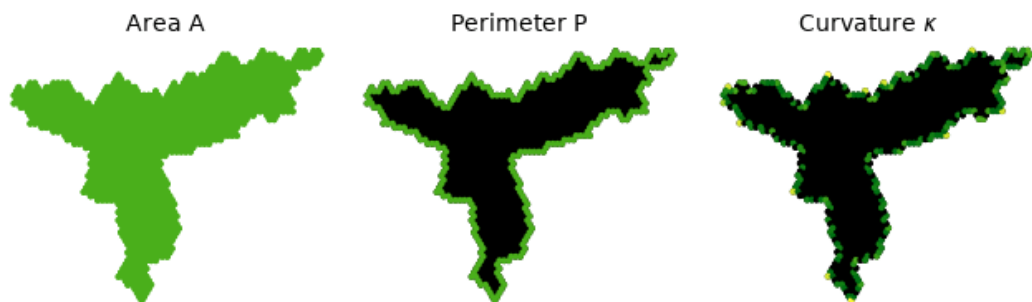
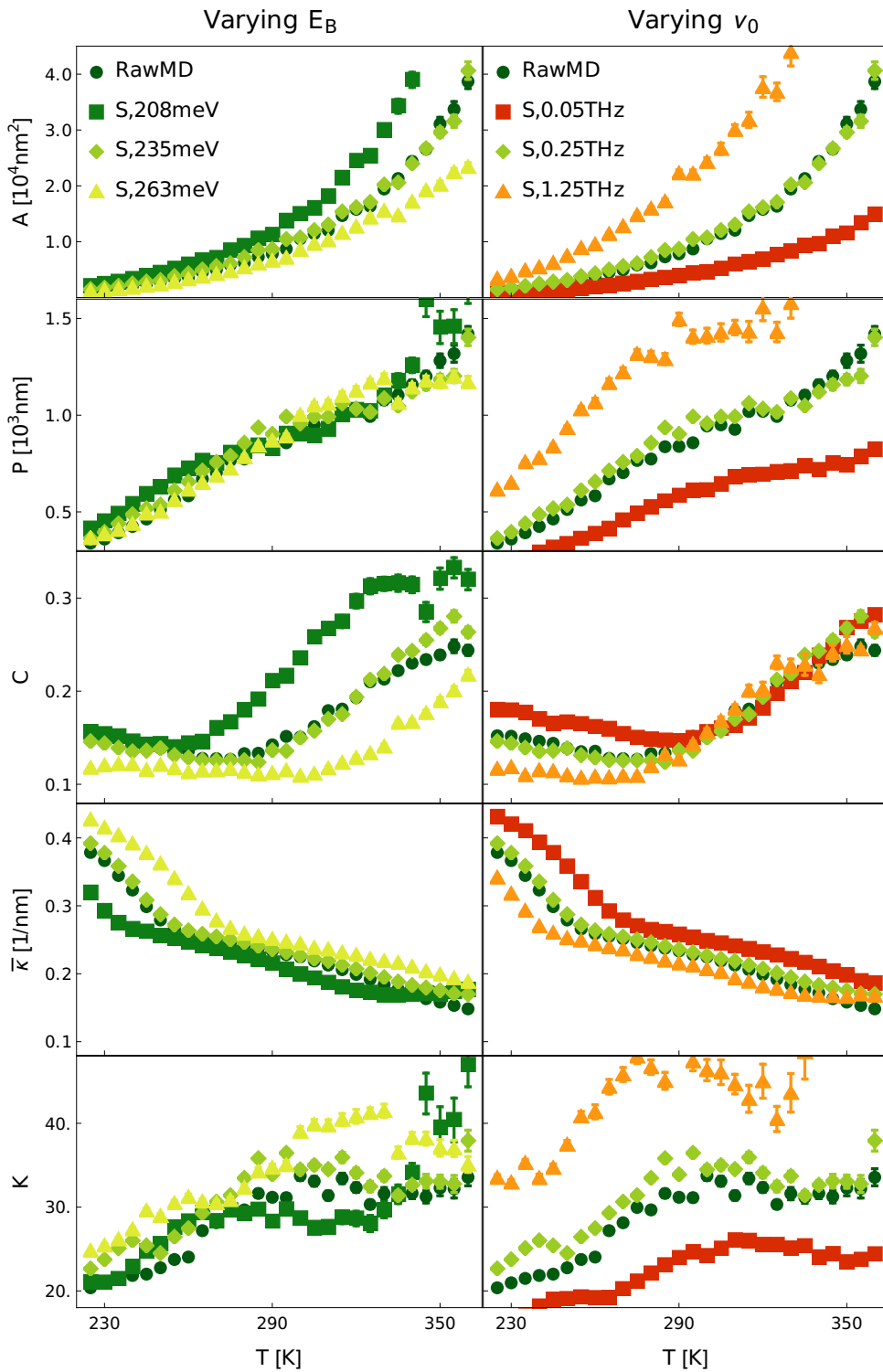


Figure 38: The three measured shape descriptors.



**Figure 39:** Average cluster features of both rate models plotted against temperature. Left: Comparison between the RawMD model with three Simple (S) models with a constant attempt rate of  $\nu_0 = 0.25$  THz and varying  $E_B$ . Right: Comparison between the RawMD model with three Simple (S) models with a constant binding energy  $E_B = 235$  meV and varying  $\nu_0$ .

(y coordinate analogous). Plugging the derivatives of Eqs. III.13 into the formula Eq. III.12, we can calculate the curvatures  $\kappa_n$  at every border molecule position and from those the mean border curvature

$$\bar{\kappa} = \frac{1}{N_E} \sum_{n=1}^{N_E} \kappa_n \quad (\text{III.14})$$

In addition to those three shape descriptors, we are also calculating two dimensionless quantities, the circularity  $C$  and the dimensionless curvature  $K$ ,

$$C = \frac{2\pi A}{P^2} \quad K = \frac{\bar{\kappa} P}{2\pi}, \quad (\text{III.15})$$

both defined to return unity for circles of any size.

The average values of the shape descriptors are plotted against temperature in Fig. 39, comparing the RawMD model with variations of the Simple model. We can see that a very good agreement between the two models is achieved for  $E_B = 235 \text{ meV}$  and  $\nu_0 = 0.25 \text{ THz}$  over the whole temperature range and also that deviation from those values can lead to significant differences. The bond strength  $E_B$  is close to an average of the effective bond strengths  $E_{B,i}$  of the RawMD model listed in Tab. 4 and closest to the ones of edge diffusion, which are the most important transitions when it comes to cluster relaxation and consequently cluster shapes. However, with this low  $E_B$  the cluster energy is underestimated which would lead to lower cluster densities / bigger clusters (see Sec. II.2.3). This is compensated by an attempt rate  $\nu_0$  that is slightly lower than the attempt rate for free diffusion in the RawMD model.

In the following section, we are going to use the Simple model with the determined optimal parameters to run multilayer growth simulations.

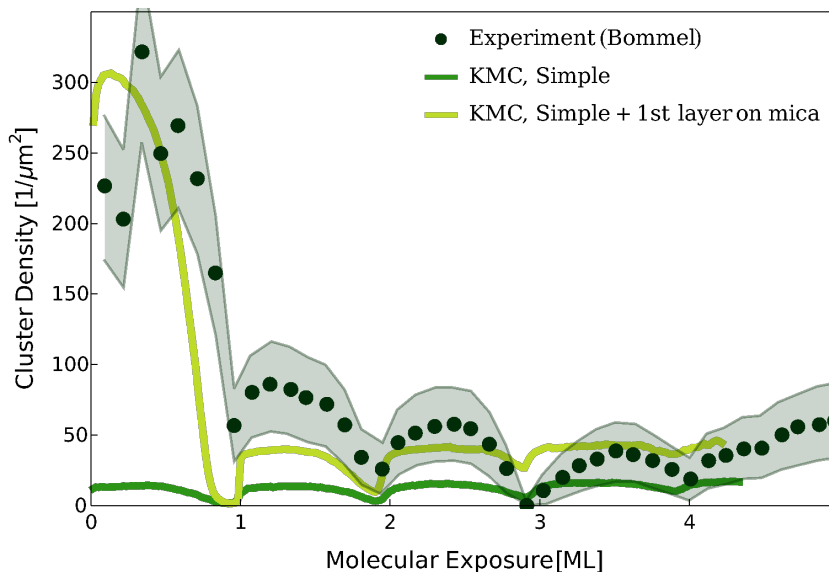
### III.2.2 MULTILAYER GROWTH

To apply our derived Simple rate model to a specific problem, we are turning to the consecutive growth of multiple  $C_{60}$  layers, for which we specifically consider the experimental data published by Bommel *et al.* in Ref. [55]. In this particular study, the time evolution of the average cluster density during deposition of  $C_{60}$  molecules on a muscovite mica ( $KAl_3Si_3O_{12}H_2$ ) substrate was measured using X-ray spectroscopy. It is plotted in Fig. 40 as a function of mean film height in units of monolayers (ML, molecular exposure =  $F \cdot \text{time} / \bar{n}$  with  $\bar{n}$  the density of a full monolayer). Using the same molecular flux and temperature as in this experiment ( $F \approx 0.1 \text{ ML/min}$ ,  $T = 333 \text{ K}$ ) we have measured the cluster density time evolution of the Simple model and plotted it in Fig. 40 (green line) in

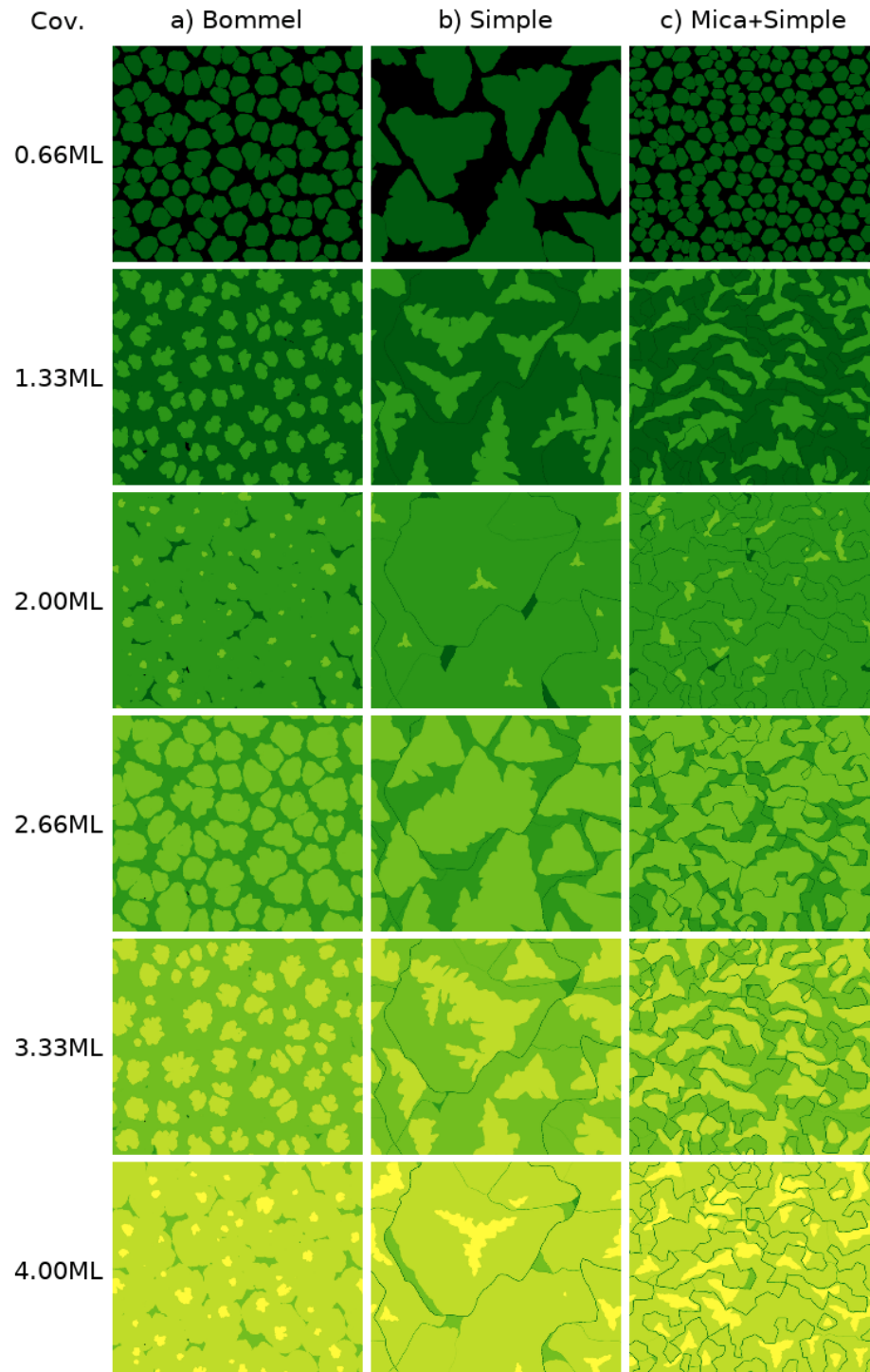
comparison to the experimental results. While the simulations naturally recover the period of the oscillations, the Simple model yields consistently smaller cluster densities, with a discrepancy exceeding an order of magnitude in both the first and second layer. The discrepancy in the first layer is to be expected as the diffusion process of  $C_{60}$  on mica is different than on  $C_{60}(111)$ , however, the behaviour of the second and higher layers warrants further investigation.

The authors of this study also reproduced parts of their data (second layer onwards) using a basic KMC simulation (no sublattices, no distinction between A and B step edge diffusion) with a top-down constructed rate model (which we refer to as "Bommel"), the parameters of which are at variance with our results and other studies in the literature:  $\Delta E_{FD,Bommel} = 540$  meV for the free diffusion barrier and  $E_{B,Bommel} = 130$  meV for the bond strength versus our values in the Simple model,  $\Delta E_{FD} = 192$  meV and  $E_B = 235$  meV. The reasons for these large differences can be seen in the snapshots of the corresponding KMC simulations in Fig. 41 [subfigures (a) are a reproduction of their model from our side]. The large diffusion barrier of  $\Delta E_{FD,Bommel} = 540$  meV was chosen to reproduce the higher cluster densities that they observed from the second layer onwards, whereas our Simple model grows much bigger clusters in lower densities. However, with this high diffusion barrier (and because they implemented a different bond counting approach), a low bond strength

**Figure 40:** Cluster density measured at  $T = 333$  K as a function of molecular exposure. Plotted are experimental results from Ref. [55] (dark green circles) in comparison to results from KMC simulations employing the pure Simple model (green line), representing the homoepitaxial growth of  $C_{60}$  on a clean  $C_{60}(111)$  surface. Also plotted are the results of the Simple model with modified diffusion rates on the first layer to reproduce the experimental first layer cluster densities (bright green line), representing heteroepitaxy of  $C_{60}$  on mica.







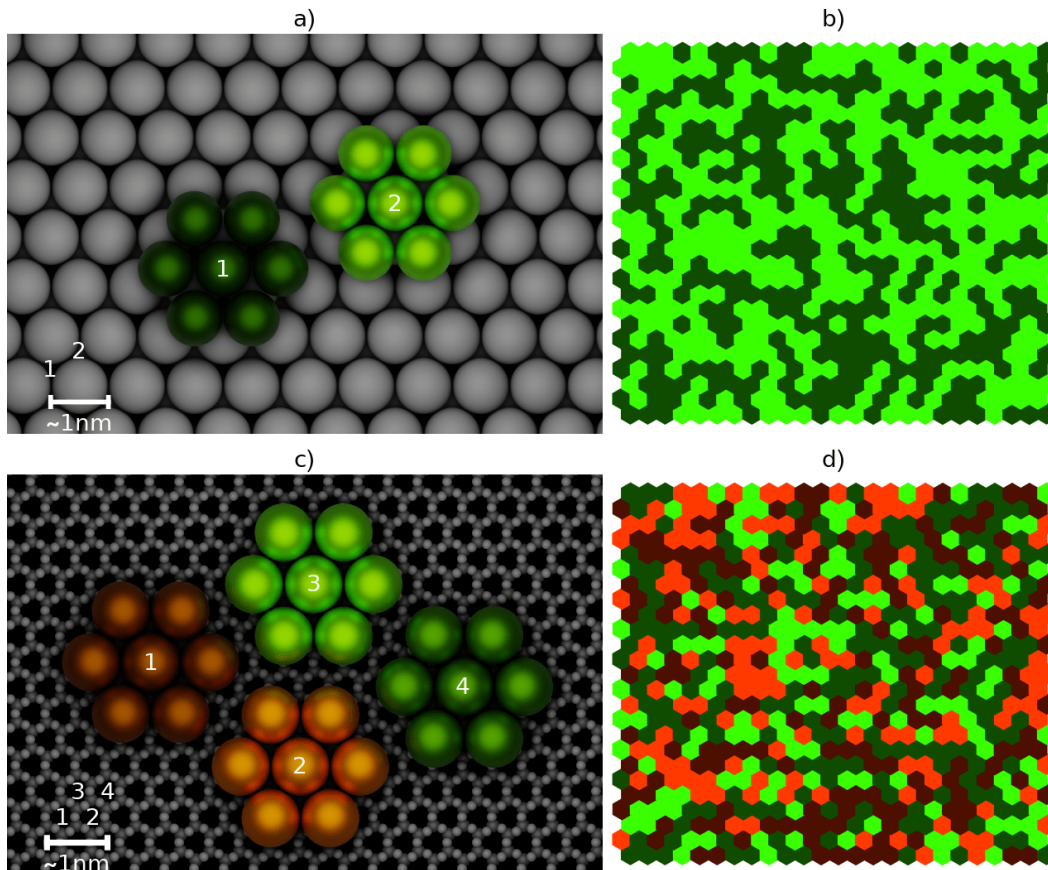
**Figure 41:** Snapshots taken at different stages of the multilayer growth simulations with a temperature of  $T = 333$  K and a molecular flux of  $F = 0.1$  ML/min. (a) Rate model as used by Bommel in Ref. [55]. (b) "Simple" rate model as derived in the previous section [notice the grain boundaries]. (c) Using the Simple rate model only for the second and higher layers. In the first layer, the diffusion barrier and attempt rate have been adjusted to reproduce the experimentally observed first layer cluster densities of  $C_{60}$  on mica given in Ref. [144].

( $E_{B,Bommel} = 130 \text{ meV}$ ) is needed to make the clusters grow into compact shapes, which may have been the goal of this parameter tuning. In contrast, the clusters in our Simple model grow into star-shaped morphologies just as we have shown earlier.

### III.2.2.1 Grain Boundaries

While the obtained diffusion barrier of  $\Delta E_{FD,Bommel} = 540 \text{ meV}$  was presented as a free diffusion barrier for the diffusion process of  $C_{60}$  on a defect-free  $C_{60}(111)$  surface, we are proposing that it is actually an effective diffusion barrier for  $C_{60}$  on a defective  $C_{60}$  surface grown on mica. As defects, we are considering "grain boundaries" that can evolve during the coalescence of a monolayer. In our KMC simulations, grain boundaries evolve when two clusters that are initialized on different sublattices (of which we have two possible

**Figure 42:** Grain boundaries and the evolution of separated domains. (a,c) Visualization of grain boundaries as a result of multiple possible cluster species (clusters evolving on different sublattices) on (a)  $C_{60}(111)$  and (c) mica. The white numerals denote the sublattice positions. (b,d) Example result of domains separated by grain boundaries for (b) two and (d) four possible sublattices. Generated by placing  $25 \times 25$  hexagons with random colors, simulating homogeneous growth of clusters with fixed spacings. [compare subfigure (b) with Fig. 41c]



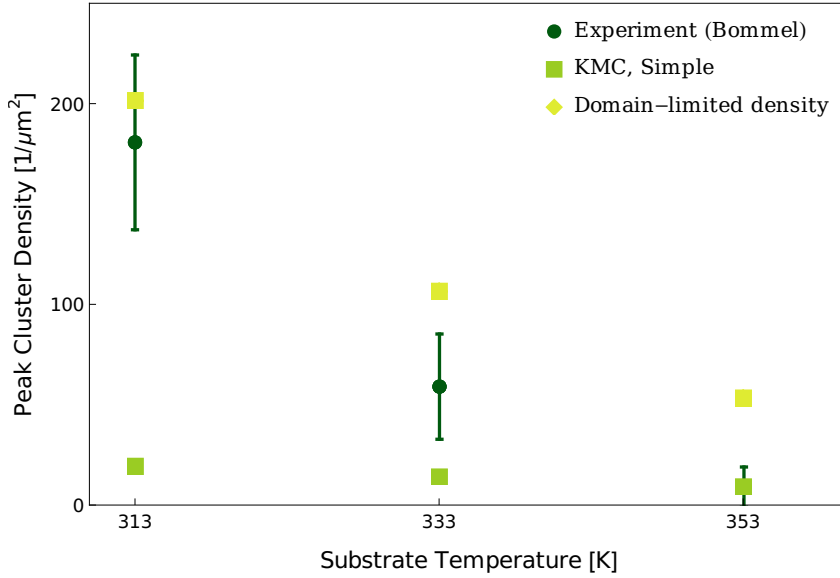
choices) of the honeycomb lattice are unable to cleanly coalesce, leaving a small space between the clusters that cannot be filled by other  $C_{60}$  molecules [Fig. 42a]. While in reality particles can take positions on top of the grain boundaries, these positions are off-lattice for the KMC simulation, so particles in the simulation are not able to cross grain boundaries. While the impossibility of crossing overestimates the diffusion hindering effect of grain boundaries, the two sublattices that we have implemented underestimate the number of grain boundaries that can evolve during the growth of  $C_{60}$  on mica. The mica surface has a lattice constant of roughly  $0.5\text{ nm}$  [145–149], resulting in four possible sublattices on which the  $C_{60}$  molecules can nucleate into clusters [Fig. 42c], all being unable to cleanly coalesce with clusters nucleated on one of the other sublattices. Therefore,  $C_{60}$  grown on mica will form many domains separated by grain boundaries [Fig. 42d], strongly hindering the diffusion of particles on the second layer.

To test the effect of grain boundary limited diffusion in our simulation, we modify the Simple model with a different behaviour on the first layer, meant to reproduce the observed first layer results of Ref. [55]. To that end we remove the distinction between A and B step edge diffusion to get compact cluster growth, we choose an ascension barrier large enough to ensure layer-by-layer growth,  $\Delta E_{\text{Asc,mica}} = 1150\text{ meV}$  and most importantly, we have determined a diffusion barrier and attempt rate of

$$\Delta E_{\text{FD,mica}} = 655\text{ meV} \quad \nu_{0,\text{mica}} = 4 \times 10^{15}\text{ Hz} \quad (\text{III.16})$$

by tuning these parameters to reproduce the experimental first layer cluster densities at temperatures  $T \in \{313, 333, 353\}\text{K}$  (data from Ref. [144]). The results of the KMC simulation with this altered first-layer behaviour are shown in Fig. 41c and 40 (bright green line). It can clearly be seen how the high first-layer cluster density has a strong impact on the growth of the subsequent layers because of the diffusion-hindering grain boundaries. Consequently, the time evolution of the cluster densities is now in better agreement with the experimental data at  $T = 333\text{ K}$  not only for the first but also for the higher layers [Fig. 40 bright green line].

While the cluster densities obtained from our Simple rate model (without the mica adjustments) can serve as a lower bound that the experiment should hit if the hindering effect of grain boundaries is negligible, we can calculate an upper bound for the case where the grain boundaries formed by the four cluster species are not crossable. To do this we take the simplified growth model shown in Fig. 42d (hexagonal clusters homogeneously growing with fixed intercluster



**Figure 43:** Peak cluster densities of the second  $C_{60}$  layer at the three experimentally observed substrate temperatures [144] (dark green discs) in comparison to the results of KMC simulations with the Simple rate model (green squares). Also shown is the upper bound for domain-limited diffusion given by Eq. III.17 (light green diamonds).

spacings) and calculate how many domains  $n_{\text{domains}}$  are forming on average from the coalescence of  $n_1$  first layer clusters. We find

$$n_{\text{domains}} \approx 0.38 \cdot n_1 \quad (\text{III.17})$$

as an upper bound for the second layer cluster densities. In Fig. 43 we plot the peak second layer densities of the experiment and our Simple model (lower bound) as well as the domain-limited upper bound of Eq. III.17 for comparison. Interestingly, the experiment perfectly matches the second layer peak density of our Simple model for the highest temperature of  $T = 353$  K while it strongly deviates for the other two temperatures and comes close to the domain-limited upper bound at  $T = 313$  K. This suggests that at low temperatures, particle diffusion and cluster growth on top of the first layer are indeed limited to domains separated by grain boundaries, while at high temperatures the  $C_{60}$  diffusion can overcome the barriers of the grain boundaries, recovering the behaviour of a clean  $C_{60}(111)$  substrate.

### III.3 SUMMARY

In this chapter, we have studied the epitaxial film growth of  $C_{60}$  on  $C_{60}(111)$ . We have employed MD simulations to systematically determine energy barriers

and attempt rates (Tab. 1, 2, 4 and 5) for the elementary diffusive processes from trajectories starting with specific initial configurations. We have shown that a rate model that is naively constructed to reproduce the MD simulation results ("RawMD" model) does not obey detailed balance by measuring the entropy production along KMC simulation trajectories. To obtain a thermodynamically consistent rate model, we constructed the "Simple" model by constraining the attempt rates  $\nu_{0,i}$  and effective bond strengths  $E_{B,i}$  to single values  $\nu_0$  and  $E_B$  and have determined the values

$$E_B = 235 \text{ meV} \quad \nu_0 = 0.25 \text{ THz} \quad (\text{III.18})$$

as optimal choices for the Simple model by matching the cluster morphologies of the Simple and RawMD model.

Using this Simple rate model, we have then investigated the multilayer growth of  $C_{60}$  in KMC simulations. We compared the time evolution of cluster densities during multilayer growth in our simulation with the experimental data from Refs. [55, 144] and found a significant discrepancy. We have then shown that this discrepancy can be explained by the evolution of grain boundaries during coalescence of the first  $C_{60}$  layer on mica, which hinders the diffusion and cluster growth process for the consecutive layers. With this, we have also found a reason why the parameters obtained for the KMC rate model of Ref. [55] are in strong disagreement with our values and other estimates in the literature.

The  $C_{60}$  on  $C_{60}(111)$  system simulated via the coarse-grained Girifalco potential was an ideal system for a bottom-up rate modelling from MD simulations as we have demonstrated here. The MD simulations achieved time scales long enough to observe all of the relevant transitions with little computational effort, enabling the gathering of an abundant amount of statistics on every single transition type. However, it is not necessary to gather this much data on every transition of interest to derive a reasonable rate model for KMC simulations. We will demonstrate this in the next chapter where we are going to apply this approach to the  $C_{60}$  on  $\text{CaF}_2(111)$  system.



# IV

## $C_{60}$ ON $CaF_2(111)$

### CONTENTS

---

iV.1	MD Simulations . . . . .	71
iV.1.1	Modelling and Interaction Potentials . . . . .	71
iV.1.2	Free Diffusion . . . . .	75
iV.1.3	Transition Rate Modelling . . . . .	86
iV.1.4	Edge Diffusion . . . . .	89
iV.1.5	Summary . . . . .	96
iV.2	KMC Simulations . . . . .	98
iV.2.1	Cluster Densities . . . . .	99
iV.2.2	Cluster Morphologies . . . . .	104

---

In this chapter, we are going to apply the bottom-up modelling approach of chapter III to  $C_{60}$  on  $CaF_2(111)$  diffusion processes to combine the results with the "Simple" model of  $C_{60}$  on  $C_{60}(111)$  diffusion to obtain a complete rate model for  $C_{60}$  on  $CaF_2(111)$  epitaxy. The contents of the first half of this chapter (MD simulations and rate modelling) are largely based on our second publication, Ref. [3].

In contrast to the  $C_{60}(111)$  substrate of the last chapter, which we could strongly coarse-grain with the Girifalco potential, the  $CaF_2(111)$  substrate is going to require a lot more computational effort as we are not able to coarse-grain it. Besides requiring the simulation of more particles to achieve the same crystal surface area,  $CaF_2(111)$  is composed of  $Ca^{2+}$  and  $F^-$  ions, requiring the implementation of computationally expensive long-ranged electrostatic interactions. This significantly reduces our ability to gather statistics on the various transition types, which is why we are limiting the scope of the following MD simulations to initial configurations with a maximum of two initial neighbours.

While the  $CaF_2(111)$  substrate cannot be coarse-grained, its interactions with the deposited  $C_{60}$  molecules can. Using a similar approach as for the derivation of the Girifalco potential, we can calculate coarse-grained  $C_{60}$ -Ca and  $C_{60}$ -F interaction potentials from atomistic C-Ca and C-F potentials by integrating out the rotational degrees of freedom of the  $C_{60}$  molecule. As the force calculations of the  $CaF_2(111)$  itself are going to require most of the computation time anyway, coarse-graining the  $C_{60}$  molecule does not yield a significant computational gain in this case. However, it is an interesting question what kind of impact a coarse-graining like this can have on the simulation of diffusion processes on a fine-grained surface like  $CaF_2(111)$  (especially in a low-temperature regime where the coarse-grained Girifalco potential is suspected to be inaccurate). We are therefore adding a comparative study of an atomistic rigid-body representation and a coarse-grained central body representation of the  $C_{60}$  molecule diffusion on  $CaF_2(111)$ . The dynamics of the two models are going to be compared in Sec. IV.1.2 with a detailed analysis of the free diffusion process and we are going to construct two separate rate models for use and comparison in KMC simulations based on coarse-grained and atomistic MD simulations.

Another difference of this system is that we do not have any established  $C_{60}$ - $CaF_2(111)$  interaction potentials, so we have to come up with them ourselves. In defining those interaction parameters we are going to leave a free parameter  $\epsilon_F$  that allows us to tune the  $C_{60}$ - $CaF_2(111)$  interaction strength in the MD simulations. This parameter will be transferred into the resulting KMC rate models as a free parameter that allows for interpolation between the observed MD simulation rates. We can then tune  $\epsilon_F$  in KMC simulations to reproduce experimentally observed cluster properties.



## IV.1 MD SIMULATIONS

### IV.1.1 MODELLING AND INTERACTION POTENTIALS

#### *CaF<sub>2</sub> Interaction Potentials*

To model the CaF<sub>2</sub>(111) substrate atoms, we follow the example of Gillan's studies on CaF<sub>2</sub> [150]. The electrostatic interaction between the Ca<sup>2+</sup> and F<sup>-</sup> ions (represented by point particles) is superimposed with a Born-Mayer-Huggins style potential,

$$\phi_{ij}^S(r) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r} + A_{ij} e^{-r/\rho_{ij}} - \frac{C_{ij}}{r^6}. \quad (\text{IV.1})$$

The electrostatic interaction is determined by the charges of the ions ( $q_{\text{Ca}} = 2e$  and  $q_{\text{F}} = -e$ ) and the parameters for the Van der Waals and repulsive interactions are given in Tab. 6. A detailed discussion of these interaction parameters can be found in Ref. [150].

**Table 6:** Parameters taken from Ref. [150] for the repulsive and Van der Waals interactions of  $\phi_{ij}^S$  [Eq. (IV.1)].

	$A_{ij}[\text{eV}]$	$\rho_{ij}[\text{\AA}]$	$C_{ij}[\text{eV}\text{\AA}^6]$
F-F	1808.0	0.293	109.1
Ca-F	674.3	0.336	0
Ca-Ca	0	—	0

In the MD simulation setups, the CaF<sub>2</sub>(111) substrate consists of four layers of CaF<sub>2</sub> arranged in a fluorite structure and is terminated by a layer of fluoride in the (111) plane (as shown in Sec. II.1.3 Fig. 5). The substrate atoms are initialized on the corresponding lattice positions and a Langevin thermostat and velocity rescaling are applied to equilibrate the system with random initial velocities at a given desired temperature. The bottommost layer of CaF<sub>2</sub> is fixed in place to ensure that the crystal structure stays in one position.

#### *Atomistic Fullerene Model*

In the atomistic representation of C<sub>60</sub> we neglect molecular vibrations and represent each molecule by 60 carbon atoms, grouped to form a rigid body.

The individual carbon atoms of two different  $C_{60}$  molecules interact via the Lennard-Jones potential

$$\phi_C(r) = 4\epsilon_C \left[ \left( \frac{\sigma_C}{r} \right)^{12} - \left( \frac{\sigma_C}{r} \right)^6 \right]. \quad (IV.2)$$

the parameters  $\epsilon_C$  and  $\sigma_C$  can in principle be derived from the parameters  $\alpha$  and  $\beta$  of the Girifalco potential via

$$\sigma'_C = R \left( \frac{480\beta}{\alpha} \right)^{\frac{1}{6}} \approx 3.47 \text{ \AA} \quad \epsilon'_C = \frac{2\alpha^2}{5\beta 60^2} \approx 2.86 \text{ meV}. \quad (IV.3)$$

However, using these parameters leads to a strong discrepancy between the lattice constants and cluster energies of the coarse-grained (Girifalco) and atomistic representations of  $C_{60}$  clusters on the  $CaF_2(111)$  surface, especially at low temperatures. As we are interested in analysing the effect of coarse-graining the adparticle-substrate interaction on the diffusion processes, we want to minimize the effect of the adparticle-adparticle coarse-graining on cluster properties like lattice strain, especially in the low-temperature regime where  $C_{60}$  was shown to exhibit structural phase transitions. Therefore, we are not using the parameters of eq. IV.3 and perform MD simulations of 19  $C_{60}$  molecules arranged in a hexagonal cluster on the  $CaF_2(111)$  substrate

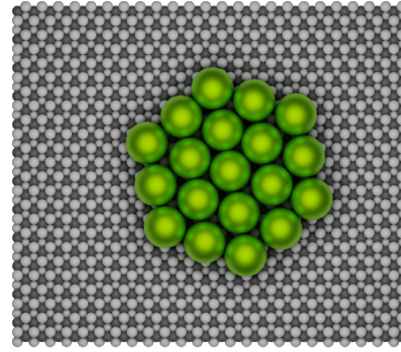


Figure 44: Configuration for the C-C interaction potential tuning.

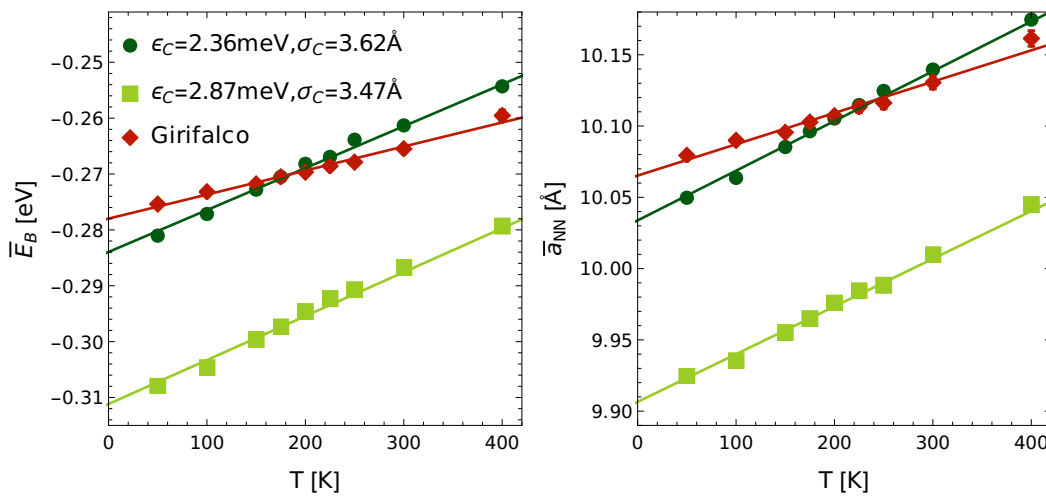


Figure 45: Average bond energy and nearest neighbour distance of  $C_{60}$  clusters on  $CaF_2(111)$  simulated with the Girifalco potential in comparison to clusters simulated using the atomistic potential with two different choices of parameters.

[Fig. 44] to match the cluster properties of the atomistic representation with the Girifalco model. The parameters  $\sigma_C$  and  $\epsilon_C$  were tuned to reproduce the average potential energy per lateral neighbour (bond energy) of  $\bar{E}_B \approx 270$  meV and an average center-to-center distance between nearest neighbours of  $\bar{a}_{NN} \approx 10.1$  Å of the Girifalco potential at  $T = 200$  K [Fig. 45]. The result of this tuning are the parameters

$$\epsilon_C = 2.36 \text{ meV} \quad \sigma_C = 3.62 \text{ \AA}. \quad (\text{IV.4})$$

For the interactions between the individual carbon atoms of the  $C_{60}$  molecules with the substrate atoms, we assume a standard Lennard-Jones C-F ,

$$\phi_F(r) = 4\epsilon_F \left[ \left( \frac{\sigma_F}{r} \right)^{12} - \left( \frac{\sigma_F}{r} \right)^6 \right], \quad (\text{IV.5})$$

and a Buckingham C-Ca interaction,

$$\phi_{Ca}(r) = A e^{-r/\rho}, \quad (\text{IV.6})$$

inspired by the interatomic interaction potentials derived in Ref. [151]. In this paper, several potentials for  $CaF_2$  with different molecules are listed. While the  $C_{60}$ - $CaF_2$  interaction was not investigated in this study, we take the listed parameters of  $CaF_2$  with other atoms of different molecules as reference points [Tab. 7]. The length scale parameters,  $\rho = 0.297$  Å and  $\sigma_F = 2.055$  Å are chosen based on the given Ca-O, Ca-Ow and CD-F potentials. The repulsive parameter  $A$  is set to be in the same range as the listed Ca-Ow, Ca-F and Ca-O potentials,  $A = 1300$  eV. The remaining attractive parameter,  $\epsilon_F$ , is the most important one, as it sets the interaction strength between carbon and fluoride atoms and consequently governs the overall adsorption energy of the  $C_{60}$  molecules on the  $CaF_2(111)$  substrate. Therefore, we are not estimating  $\epsilon_F$  and instead leave it as a free variable that we vary in a range of  $\epsilon_F \in$

**Table 7:** Compilation of some of the interaction parameters derived in Ref. [151] and the parameters that we have chosen to use for our atomistic model (bottom two lines) for eqs. IV.5 and IV.6. (O: oxygen of carbonate group, Ow: Oxygen of water, CD: carbon of methanoic acid)

ion pair	Buckingham		Lennard-Jones	
	A[eV]	$\rho$ [Å]	$\sigma$ [Å]	$\epsilon$ [mev]
Ca-O	1550	0.297	—	—
Ca-Ow	1186	0.297	—	—
Ca-F	1272	0.2997	—	—
CD-F	—	—	2.055	146
C-Ca	1300	0.297	—	—
C-F	—	—	2.055	[35, 55]

[35,55] meV, which results in a total molecule–substrate adsorption energy of 300 – 800 meV. We expect that the dewetting barrier/adsorption energy for  $C_{60}$  on  $CaF_2(111)$  is in this range since  $C_{60}$  has been observed to form two-layered clusters on  $CaF_2(111)$  at room temperature [36, 53], suggesting that the molecule–substrate interaction is somewhat weaker than the molecule–molecule interaction (again, the total adsorption energy of  $C_{60}$  on  $C_{60}(111)$  is roughly 930 meV).

### Coarse-grained Fullerene Model

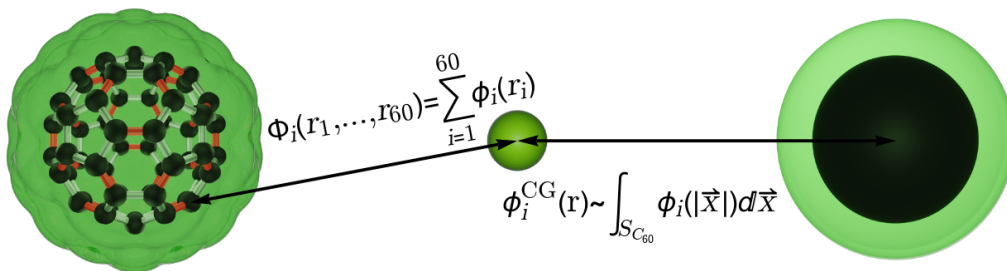
For the coarse-grained  $C_{60}$  model we employ the same Girifalco potential that we have derived in Sec. III.1.1 for the  $C_{60}$ – $C_{60}$  interaction. To calculate coarse-grained interaction potentials of a  $C_{60}$  molecule with the substrate atoms we follow a similar approach to the derivation of the Girifalco potential (and as it was previously done in other works [152–154]) and smear out the C atoms of the  $C_{60}$  molecule onto a sphere  $S_{C_{60}}$  of radius  $R = 3.55 \text{ \AA}$  over which we integrate the atomistic interaction potentials  $\phi_i(r)$  ( $i \in \{Ca, F\}$ ) of eqs. IV.6 and IV.5. The formula from which the coarse-grained potentials  $\phi_i^{CG}(r)$  are obtained can be written in spherical coordinates as

$$\phi_i^{CG}(r) = -2\pi R^2 \eta \int_0^\pi d\theta \phi_i(\sqrt{R^2 + r^2 - 2rR \cos \theta}), \quad (IV.7)$$

where  $\eta = 60/(4\pi R^2)$  is the number density of carbon atoms on the sphere and  $r$  is the distance between the center of the sphere  $S_{C_{60}}$  and the substrate atom. Solving the integration for our interatomic potentials, we obtain

$$\phi_F^{CG}(r) = \frac{60}{Rr} \epsilon_F \left( \frac{\sigma_F^6}{2(R+r)^4} - \frac{\sigma_F^{12}}{5(R+r)^{10}} - \frac{\sigma_F^6}{2(R-r)^4} + \frac{\sigma_F^{12}}{5(R-r)^{10}} \right) \quad (IV.8)$$

**Figure 46:** Atomistic (left) and coarse-grained (right) interaction  $C_{60}$  molecules with a single atom.



and

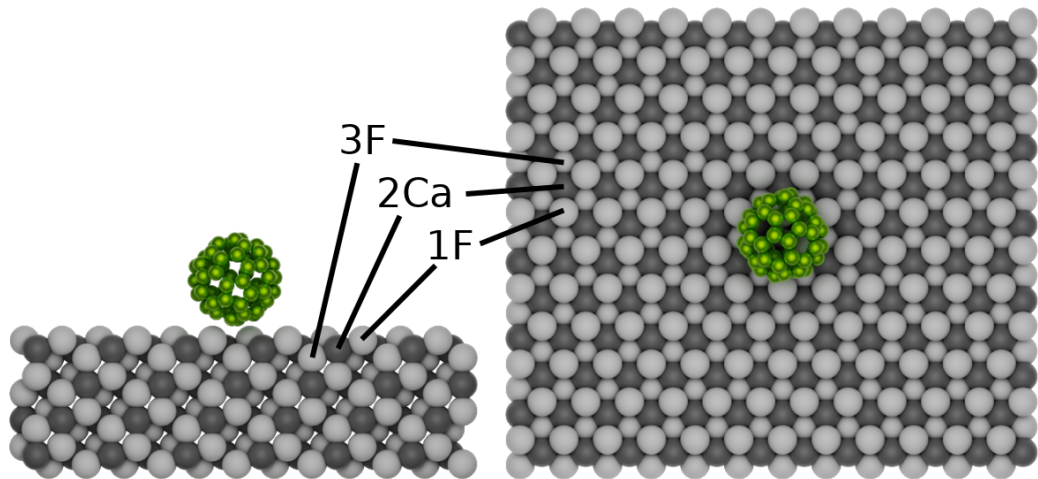
$$\phi_{\text{Ca}}^{\text{CG}}(r) = \frac{60A\rho^2}{Rr} e^{-r/\rho} \left[ \sinh\left(\frac{R}{\rho}\right) \left(1 - \frac{R}{\rho}\right) + \frac{r}{\rho} \cosh\left(\frac{R}{\rho}\right) \right]. \quad (\text{IV.9})$$

The parameters  $\epsilon_F$ ,  $\sigma_F$ ,  $A$  and  $\rho$  are assigned the same values as in the atomistic model (with  $\epsilon_F$  being a free parameter)

#### IV.1.2 FREE DIFFUSION

We start with the free diffusion process, where a single  $\text{C}_{60}$  molecule is diffusing on a clean  $\text{CaF}_2(111)$  surface at a specific substrate temperature. The simulations are again set up using the LAMMPS simulation package [118]. The simulation box is sized approximately  $40 \text{ \AA} \times 46 \text{ \AA}$  with periodic boundary conditions along the  $x$  and  $y$  direction. A single  $\text{C}_{60}$  molecule is placed in the middle of the surface at the beginning of the simulation [Fig. 47]. For the atomistic model, the interactions are implemented using the available born/coul/long and lj/cut pair potentials while for the coarse-grained model we tabulated the potentials IV.8 and IV.9 to be usable in LAMMPS. After equilibrating the substrate to the desired temperature, a 50 ns trajectory of the  $\text{C}_{60}$  molecules is recorded for later analysis. For each temperature ( $T \in [40, 450] \text{ K}$ , covering low to medium-high temperatures) and interaction strength ( $\epsilon_F = \{35, 40, 45, 50, 55\} \text{ meV}$ ) ten of these trajectories are run, totalling 500 ns of diffusion per set of parameters.

**Figure 47:** Side (left) and top view (right) of the MD simulation setup for the free diffusion process of  $\text{C}_{60}$  on  $\text{CaF}_2(111)$ . Carbon atoms are shown in green, calcium atoms in grey and fluoride atoms in light grey



#### IV.1.2.1 Minimum Energy Paths

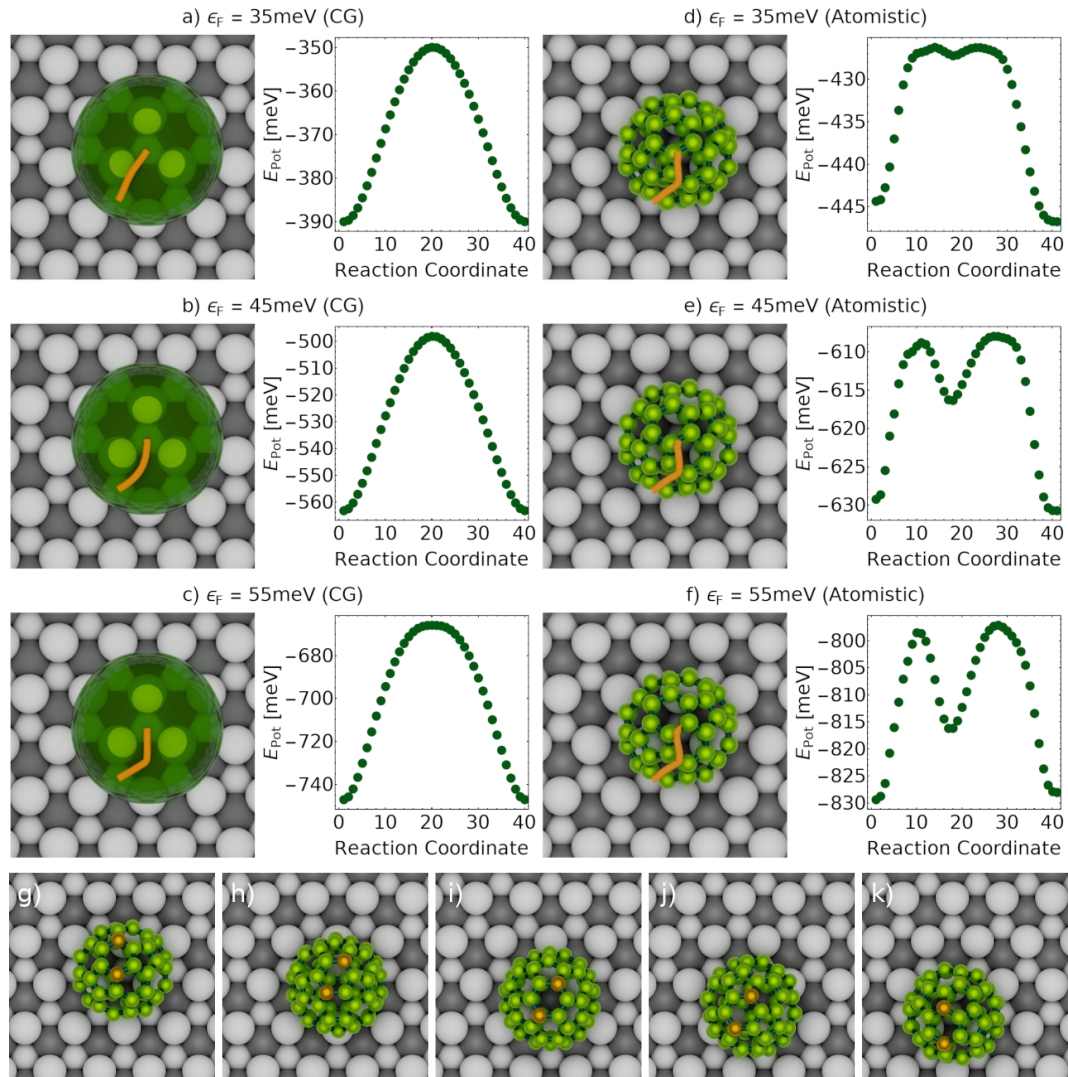
Before going over the actual MD simulation results, we are taking a closer look at the minimum energy paths (MEPs) for the surface diffusion of a  $C_{60}$  molecule in both the atomistic and coarse-grained model.

From the MD simulation trajectories, we know that in both models the  $C_{60}$  molecules preferably occupy positions above third-layer fluoride atoms (position 3F in Fig. 47), which we identify as lattice sites of the free diffusion process. To determine the MEPs from a lattice site to a neighbouring one, we have implemented the drag and NEB method (see Sec. II.3.4) in Mathematica [155]. For simplicity, we leave the  $CaF_2(111)$  substrate stationary and only optimize the position/rotation of the  $C_{60}$  molecule along the MEPs. In the coarse-grained case, we choose a straight line connecting the two lattice sites as an initial path for the NEB method. In the atomistic model, the choice of an initial path is not as trivial because of the rotational degrees of freedom. We therefore first generate an MEP for the atomistic model using the drag method and then further optimize this MEP with the NEB method. Illustrations of the resulting MEPs and their corresponding potential energy plots are shown in Fig. 48(a-f) and intermediate snapshots from the MEP of the atomistic model with  $\epsilon_F = 45$  meV are shown in Fig. 48(g-k). The energy minima ( $E_{\min}$ ), maxima ( $E_{\max}$ ) and barriers ( $\Delta E_{FD}$ ) derived from the potential energy plots are compiled in Tab. 8.

Major differences between the two models can be spotted in these MEP results. Firstly, the bond between the atomistic molecule and the substrate is 10–20% stronger, expressed by a larger value of  $E_{\min}$ . This difference is expected due to the atomistic model's ability to align the orientation of the molecule with respect to the substrate atom positions. Secondly, as the atomistic molecule can also adjust its orientation during the diffusive transition, it has a significantly lower diffusion barrier than the coarse-grained model. The rotational degrees of freedom also cause the atomistic model to have an additional local energy minimum between the two lattice sites on top of the surface calcium by aligning the molecule face-down [Fig. 48i]. This additional energy minimum is quite weak for  $\epsilon_F = 35$  meV but gets stronger with increasing  $\epsilon_F$  (see Fig. 48(d-f)),

**Table 8:** MEP analysis results. All values are given in meV

$\epsilon_F$	Atomistic			Coarse-Grained		
	$E_{\min}$	$E_{\max}$	$\Delta E_D$	$E_{\min}$	$E_{\max}$	$\Delta E_{FD}$
35	-447	-426	20.5	-390	-350	40.1
40	-537	-516	21.2	-475	-421	54.0
45	-631	-608	22.8	-563	-498	65.1
50	-728	-702	25.8	-654	-580	73.8
55	-829	-797	32.3	-747	-666	81.2



**Figure 48:** Minimum energy paths (MEPs) for surface diffusion in (a-c) the coarse-grained and (d-f) the atomistic model for three values of  $\epsilon_F$ . Subfigures (a-f) are split into an initial configuration (left) with the center-of-mass trajectory of the MEP shown in orange and a plot of the potential energy along the MEP (right). (g-k) Snapshots of the atomistic MEP at  $\epsilon_F = 45$  meV with two atoms of the  $C_{60}$  molecule highlighted in orange to track the orientation of the molecule.

as the C-F attraction overwhelms the repulsion of the C-Ca interaction. With the presence of the meta-stable "intermediate lattice site" on top of the surface calcium, the atomistic model tends to always take the path over the calcium location. In contrast, the coarse-grained model follows a more direct path between the lattice sites at  $\epsilon_F = 35$  meV and only passes over the calcium location at stronger values of  $\epsilon_F$  (without the presence of an actual energy minimum).

#### IV.1.2.2 Energy Landscapes

The MEPs and their corresponding potential energy plots of the last section give an idea of how the potential energy landscapes of our models look like. In this section, we extract the free energy landscape from the MD trajectories at non-zero temperatures to see if they match the expectations set by the MEPs. As the  $CaF_2(111)$  surface can be seen as a repetition of a simple unit cell (shown in Fig. 49) in  $x$  and  $y$  direction, we project the center-of-mass trajectories into a single unit cell to best make use of the available data. Furthermore, we assume that the positional probability distribution can be written as

$$p(x, y; T) = A \cdot \exp\left(\frac{-F(x, y; T)}{k_B T}\right), \quad (IV.10)$$

with a normalization constant  $A$  and a free energy function  $F(x, y; T)$ . As we are mainly interested in energy barriers, we solve Eq. IV.10 for  $F$  and write it as a difference:

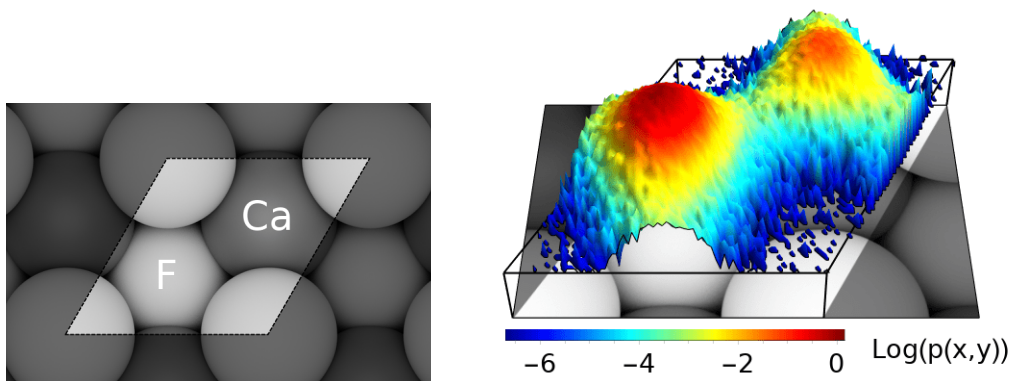
$$F(x, y; T) = k_B T (\log(A) - \log(p(x, y; T))) \quad (IV.11)$$

$$\Delta F(x_1, y_1, x_2, y_2; T) = k_B T [\log(p(x_1, y_1; T)) - \log(p(x_2, y_2; T))]. \quad (IV.12)$$

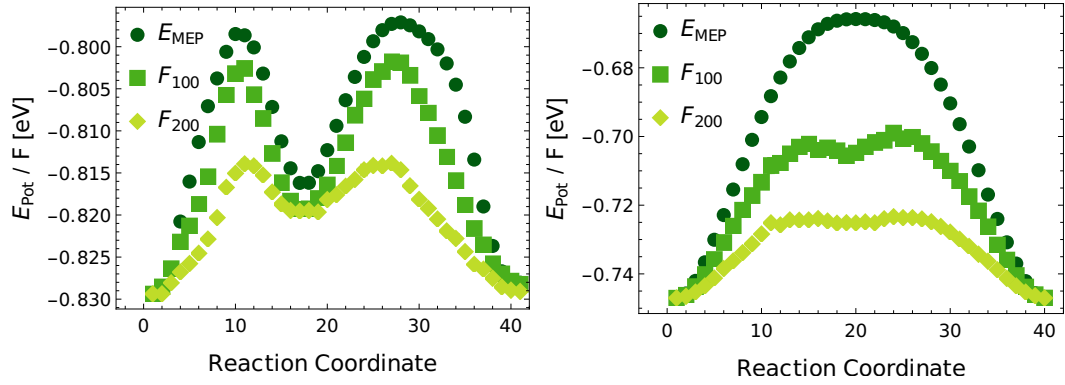
With Eq. IV.12 we can now reconstruct a free energy landscape from the positional probability distribution except for a constant offset. Setting this offset to match the minimum energy of the MEPs, we can compare the potential energy to the free energy landscape along the MEPs.

In Fig. 50 we can already see that the overall shapes of the free energies are in quite good agreement with the shapes of the potential energies along the MEPs. Also, in Fig. 51(left) we can see that the MEPs nicely align with the obtained free energy landscapes. Especially the presence of a second local

**Figure 49:** Left:  $CaF_2(111)$  surface unit cell. Right: Visualization of the logarithmic positional probability distribution of a  $C_{60}$  molecule inside a  $CaF_2(111)$  unit cell. The shown distribution is taken from the MD trajectories of the atomistic model with  $\epsilon_F = 55$  meV at  $T = 100$  K.



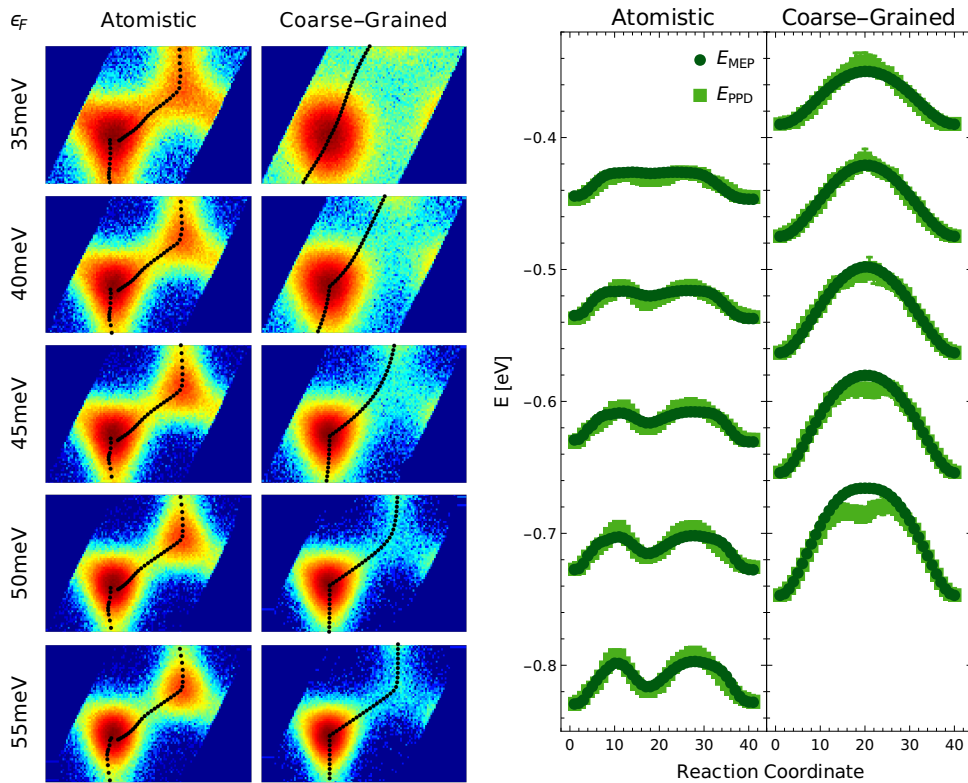




**Figure 50:** Potential energy  $E_{\text{MEP}}$  (as calculated in the previous section) in comparison to the free energies  $F_{100}/F_{200}$  (at  $T = 100$  K and  $T = 200$  K respectively) along the MEP at  $\epsilon_F = 55$  meV (Left: atomistic. Right: coarse-grained).

energy minimum in the unit cell above the calcium for the atomistic model as well as its absence in the coarse-grained model are very nicely reproduced in these density plots. However, as the free energy strongly depends on the

**Figure 51:** Left: Density plots of  $\ln(p(x, y; T))$  at  $T = 100$  K for several values of  $\epsilon_F$  and for both models with the corresponding MEPs overlaid as black dots. Right: Comparison of the potential energies  $E_{\text{MEP}}$  along the MEPs (from the previous section) to the potential energies  $E_{\text{PPD}}$  along the MEPs (obtained via extrapolation of the free energy MD data to 0 K). The binding strength  $\epsilon_F$  increases from 35 meV for the topmost plot to 55 meV for the bottommost plot in steps of 5 meV.



temperature of the system, we can only expect agreement with the previously determined potential energies along the MEPs at  $T = 0$  K. To extrapolate from our data to  $T = 0$  K we use the definition

$$F(x, y; T) = E(x, y) - TS(x, y), \quad (\text{IV.13})$$

where we have assumed that there is only a linear dependence between  $F$  and  $T$  and that the positional energy  $E(x, y)$  and entropy  $S(x, y)$  are independent of temperature. Taking our free energy measurements at various temperatures, we can obtain  $E(x, y)$  and  $S(x, y)$  for any point in the unit cell via a linear fit. Doing this for the coordinates of the MEPs yields the results shown in Fig. 51 (right), which are plotted in direct comparison to the previously determined potential energies along the MEPs. We can see that the results we obtained with these two different approaches are in very good agreement with each other.

#### IV.1.2.3 Diffusion Coefficients

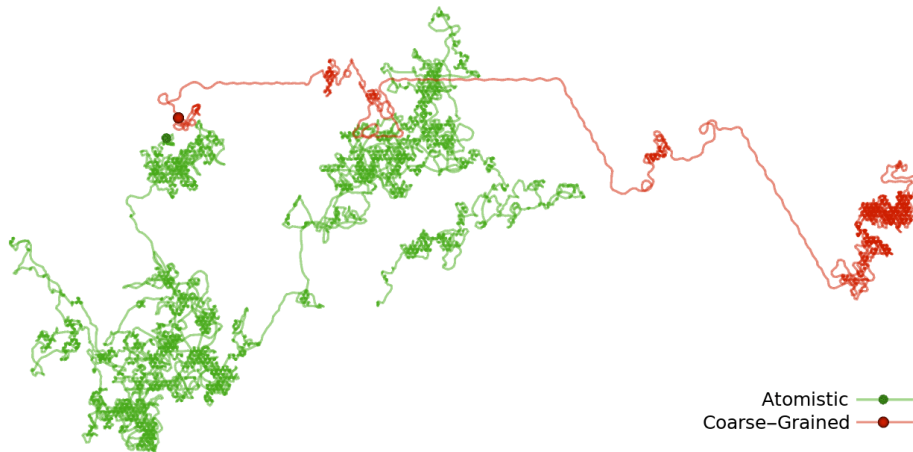
For a final comparing analysis between the two models, we are going to apply the common approach of evaluating the mean-squared displacement (MSD, see sec. II.3.5) of the MD trajectories. To derive diffusion coefficients  $D(T)$ , diffusion barriers  $\Delta E_{\text{FD}}$  and attempt rates  $\nu_{0,\text{FD}}$  from the MSDs, we employ the formulas

$$\text{MSD}(\Delta t) = \langle [\mathbf{x}(t + \Delta t) - \mathbf{x}(t)]^2 \rangle = 4D(T)\Delta t, \quad (\text{IV.14})$$

$$D(T) = \frac{1}{4}\nu_{0,\text{FD}}\langle l^2 \rangle \exp\left(-\frac{\Delta E_{\text{FD}}}{k_{\text{B}}T}\right), \quad (\text{IV.15})$$

where we set the mean-squared jump length to  $\langle l^2 \rangle = 0.386^2 \text{nm}^2 = 0.149 \text{nm}^2$  to obtain effective attempt rates  $\nu_{0,\text{FD}}$  under the assumption of single jumps

**Figure 52:** Example free diffusion trajectories of the atomistic (green, measured at  $T = 250$  K) and coarse-grained model (red, measured at  $T = 150$  K).

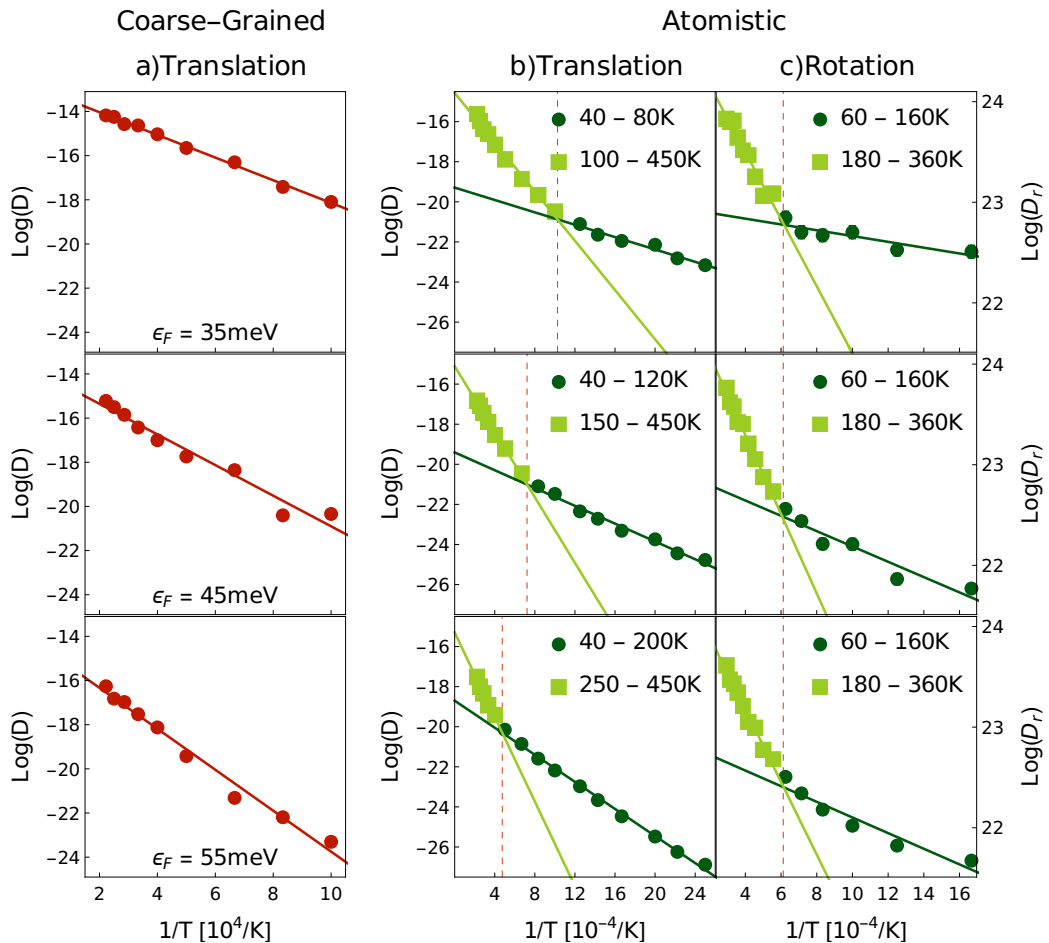


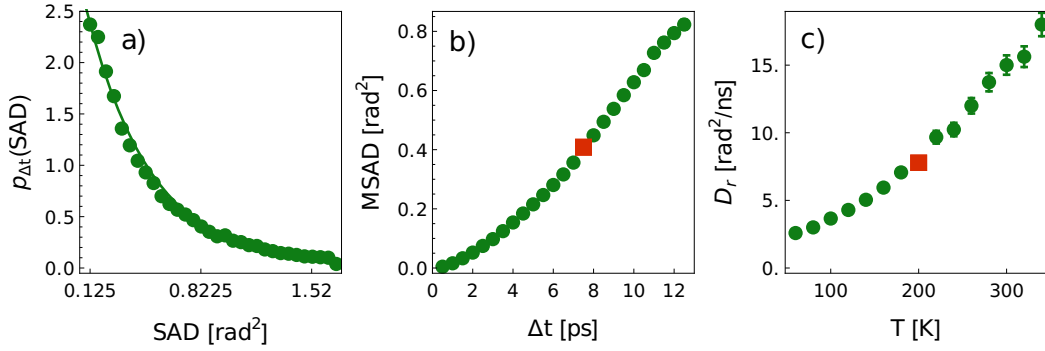
between neighbouring lattice sites. However, for the discussion of the resulting attempt rates, we will have to keep in mind that this assumption is not accurate in our case since the MD trajectories show jumps across multiple lattice sites (especially at higher temperatures). In the coarse-grained model, we can even observe occasional long diffusive jumps without a change in direction [Fig. 52], similar to the behaviour described by Lévy flights [156, 157].

The resulting Arrhenius plots are shown in Fig. 53. While the coarse-grained model turns out to follow an Arrhenius law across the board (the data points in Fig. 53a fall onto lines), the atomistic model exhibits a crossover in diffusive behaviour at crossover temperatures  $T_{\text{Cross}} \in [100, 200]\text{K}$  which depend on the value of  $\epsilon_F$  [Fig. 53b].

As this crossover in the behaviour of the atomistic model hints at a relation to the structural phase transition of  $C_{60}$  [16–21], we take a look at a few additional shorter trajectories with a higher time resolution to be able to analyze the

**Figure 53:** Arrhenius plots for the obtained (a) translational diffusion coefficients of the coarse-grained model in comparison with (b,c) the translational and rotational diffusion coefficients of the atomistic model. The dashed red lines mark the crossover temperatures of the atomistic model.





**Figure 54:** Determination of the rotational diffusion coefficients from (a) distributions of squared angular displacements (SADs). (b) MSADs obtained from SAD distributions. The red data point is the result from a fit of Eq. IV.18 to the distribution shown in (a). (c) Rotational diffusion coefficients obtained from MSADs. The red data point is the result of a fit of Eq. IV.16 to the data shown in (b).

rotational diffusion of the  $C_{60}$  molecules in the atomistic model. These trajectories are obtained from simulations with  $\epsilon_F \in \{35, 45, 55\} \text{ meV}$  at temperatures  $T \in [60, 360] \text{ K}$ , covering a total of 50 ns of diffusion per set of parameters. From the trajectories, we aim to determine a rotational diffusion coefficient  $D_r(T)$  analogous to the translational case (Eq. IV.15) using a definition of the mean squared angular displacement (MSAD),

$$\text{MSAD}(\Delta t) = \langle [\theta(t + \Delta t) - \theta(t)]^2 \rangle = 2f_r D_r(T) \Delta t \quad (\text{IV.16})$$

with the number of rotational degrees of freedom  $f_r = 3$ . To calculate the MSADs, we determine the angular displacements  $[\theta(t + \Delta t) - \theta(t)]$  as the angle of optimal rotation between the conformations of the  $C_{60}$  molecule at times  $t$  and  $t + \Delta t$ . To find the optimal rotation between the two conformations, we follow a quaternion approach as described in Refs. [158, 159] (see also App. B.2). For small time intervals  $\Delta t$ , we find that the squared angular displacements ( $\text{SAD}(t, \Delta t) = [\theta(t + \Delta t) - \theta(t)]^2$ ) calculated this way are exponentially distributed [Fig. 54a],

$$p_{\Delta t}^*(\text{SAD}(t, \Delta t)) = \frac{1}{\text{MSAD}(\Delta t)} \cdot e^{-\text{SAD}(t, \Delta t)/\text{MSAD}(\Delta t)}. \quad (\text{IV.17})$$

However, with increasing  $\Delta t$ , the distribution naturally changes as the SAD of an optimal rotation can only take values  $\text{SAD}(t, \Delta t) \in [0, \pi^2]$ , effectively reflecting the exponential distribution back and forth in this interval. The resulting distribution can be written as

$$p_{\Delta t}(x) = \frac{2\lambda \cosh(\lambda x)}{1 - e^{-2\lambda\pi^2}} - \lambda e^{\lambda x}, \quad (\text{IV.18})$$

with substitutions  $\lambda = 1/\text{MSAD}(\Delta t)$  and  $\chi = \text{SAD}(t, \Delta t)$  applied for better readability. Fitting the function Eq. IV.18 to the distribution of SADs, we obtain values for  $\text{MSAD}(\Delta t)$  [Fig. 54b] from which we can then determine the rotational diffusion coefficients  $D_r$  via Eq. IV.16 [Fig. 54c]. The results are shown in the Arrhenius plots for the rotational diffusion coefficients in Fig. 53c. While the rotational diffusion also experiences a crossover, it appears consistently at a crossover temperature of around  $T_{\text{Cross,R}} = 163(1)$  K.

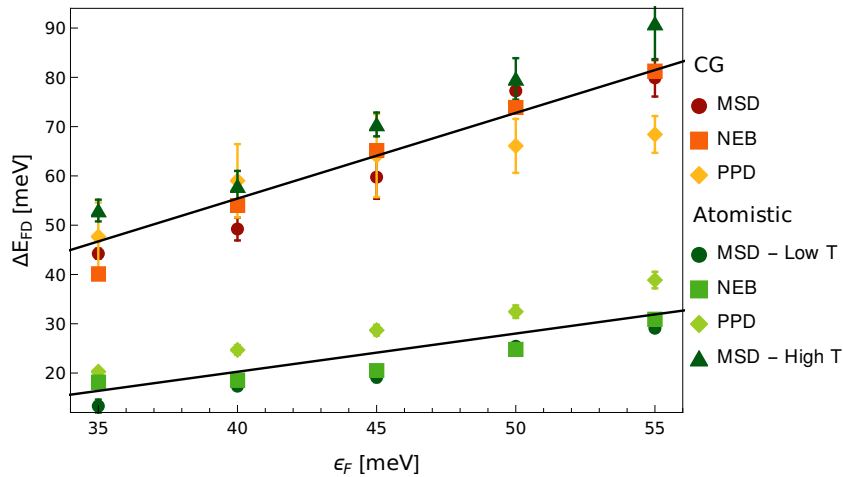
The crossover temperatures we have observed in the translational and rotational diffusion coefficients are lower than the temperatures at which structural phase transitions of  $C_{60}$  were experimentally observed in thin films [19, 20] (220 – 260 K). However, deposition experiments of  $C_{60}$  on metal-silicon surfaces have found a crossover in diffusive behaviour (or rather a crossover in island density formation behaviour) in a temperature range of 140 – 160 K [48], very compatible with our results. A similar observation of a diffusive crossover was made in MD simulations of the diffusion process of  $C_{60}$  on graphene [160] in a temperature range of 25 – 75 K. We can suspect that the crossover temperature was shifted to much lower values in this case because the graphene substrate results in much lower diffusion barriers for the  $C_{60}$  molecule than the  $\text{CaF}_2(111)$  surface. With this wide variation of experimentally and numerically observed crossover temperatures on different substrates, we can conclude that our observation of a crossover temperature that depends on  $\epsilon_F$  fits into the established work.

#### IV.1.2.4 Summary and Discussion

From the analyses of the previous sections, we have obtained several estimates for the diffusion barriers  $\Delta E_{\text{FD}}(\epsilon_F)$  for the coarse-grained and atomistic model: From the MEPs in Sec. IV.1.2.1, the energy landscapes in Sec. IV.1.2.2 and from the Arrhenius plots in Sec. IV.1.2.3. The results are summarized in Fig. 55 and can be grouped into two sets. The first set is composed of the lower energy barriers (14 – 40 meV) resulting from the low-temperature analyses of the atomistic model and is, therefore, most likely accurate for the diffusion process at low temperatures (where the rotational degrees of freedom of the  $C_{60}$  molecule play a bigger role). The second set contains the energy barriers from the coarse-grained model as well as the MSD results of the atomistic model at high temperatures, which all fall into a higher range of 40 – 90 meV. We can apply linear fits to the two sets of energy barriers to obtain models for the low temperature (LT) and high temperature (HT) regime respectively:

$$\Delta E_{\text{FD,LT}} = -10.7(67) \text{ meV} + 0.78(14) \cdot \epsilon_F \quad (\text{IV.19})$$

$$\Delta E_{\text{FD,HT}} = -14.1(82) \text{ meV} + 1.74(18) \cdot \epsilon_F. \quad (\text{IV.20})$$



**Figure 55:** Summary of free diffusion energy barriers obtained in the previous sections. The data can be grouped into two sets which are fitted separately (black lines). MSD: Diffusion barriers obtained from the translational mean-squared displacement analyses in Sec. IV.1.2.3. NEB: Diffusion barriers taken from the MEPs of the nudged elastic band method in Sec. IV.1.2.1. PPD: Energy barriers taken from the energy landscapes calculated in Sec. IV.1.2.2.

Through the observation of cluster densities in molecular beam epitaxy experiments [36], Felix Loske *et al.* determined an estimate for the diffusion barrier of  $C_{60}$  on  $CaF_2(111)$  in a temperature range of  $T \in [96, 217]$  K. The experimental estimate of  $\Delta E_{FD, Loske} = 214(16)$  meV is in strong disagreement with our results in both the high and low temperature regime. We can extrapolate our models of Eqs. IV.19 and IV.20 to get a diffusion barrier of 214 meV at values of

$$\epsilon_{F, Exp, LT} = 288(56) \text{ meV} \quad \epsilon_{F, Exp, HT} = 131(17) \text{ meV}. \quad (IV.21)$$

However, these high values of  $\epsilon_F$  are at odds with the observation of two-layered cluster growth of  $C_{60}$  on  $CaF_2(111)$  [53, 60], as the adparticle-substrate interaction would be considerably stronger than the adparticle-adparticle interaction (with ascension barriers larger than 2 eV).

A possible explanation for this discrepancy was discussed in the erratum to Ref. [36] (Ref. [2]): Impurities on the substrate surface (e.g. adsorbed air/water molecules for which cold surfaces act as a sink) could strongly hinder the diffusion process by acting as nucleation sites for the  $C_{60}$  molecules ("impurity trapping" [85]). The result would be an effective diffusion process with long jumps between those impurities, which can naturally not be compared to the diffusion of  $C_{60}$  on a clean  $CaF_2(111)$  surface as we have simulated it in this work. A different plausible explanation is the effect of polarizability of  $C_{60}$  (which we have neglected). In growth experiments of  $C_{60}$  on  $ZnPc/Ag(111)$  [134] it was shown that the polarizability of  $C_{60}$  can explain the emergence

of chain phases. We can suspect that  $C_{60}$  polarizability is also relevant on the  $\text{CaF}_2(111)$  surface, as the top fluoride layer of the substrate is negatively charged.

For the attempt rates of the free diffusion process that we can extract from the MSD analyses, we are not going into as much detail, as the discussion on the energy barriers already exposed the discrepancy. However, we are going to model the attempt rates in the following section to supply rate models on the free diffusion process for KMC simulations. We can note that the attempt rates of the atomistic model are in the order of  $10^{11}$  Hz in the low-temperature regime and  $10^{12} - 10^{13}$  Hz in the high-temperature regime, while they are in the range of  $10^{13} - 10^{14}$  Hz in the coarse-grained model. Expectedly, these are also at variance to the attempt rate obtained from the deposition experiments of Felix Loske [2] ( $\nu_{0,FD,Loske} = 3.2 \times 10^{17}$  Hz), but they are compatible with other values that have been reported for large molecules [24, 140, 141] ( $10^{10} - 10^{14}$  Hz).

Concluding this section on the free diffusion process of  $C_{60}$  on  $\text{CaF}_2(111)$ , where we have analyzed the process in detail with three different approaches, we have found that the atomistic and coarse-grained models exhibit significantly different behaviours at low temperatures. According to our mean-squared displacement analyses, the difference diminishes for higher temperatures above some crossover temperature  $T_{\text{Cross}}$ , which depends on the value of the interaction strength  $\epsilon_F$ . In line with previous expectations, our observation also suggests that the coarse-grained model should only be used in the high-temperature regime, while the atomistic model is probably the better choice for low temperatures. However, to be able to test this conclusion, we keep working with both models to measure edge diffusion transition rates in the following section. The rate models we can thereby derive for both the atomistic and coarse-grained representations of  $C_{60}$  on  $\text{CaF}_2(111)$  can then be tested in KMC simulations to see which of those models produces the better results when compared to experimentally observed cluster morphologies.

### IV.1.3 TRANSITION RATE MODELLING

With the simulations of the following section, we want to determine a rate model with a free parameter  $\epsilon_F$  for sets of transition rates observed in MD simulations. For a given transition type  $i$  from an initial state with  $n$  lateral neighbours probed at interaction strength  $\epsilon_F$  and temperature  $T$ , we estimate the transition rate via the estimator

$$k_i(\epsilon_F, n, T) = \frac{N_i(\epsilon_F, n, T)}{t_{\text{tot},i}(\epsilon_F, n, T)}, \quad (\text{IV.22})$$

where  $t_{\text{tot},i}(\epsilon_F, n, T)$  is the total time we have observed an initial state with  $n$  lateral neighbours that can go through transition type  $i$  at the specified parameters  $(\epsilon_F, T)$ .  $N_i(\epsilon_F, n, T)$  is the number of observations of transition  $i$  during this time. We can note that in Ref. [3] we have used a slightly different estimator ("unbiased") which leads to minimal differences in the results presented here. For this set of edge diffusion simulations we are considering the transition types  $i \in \{\text{ED-A, ED-B, Asc, Diss}\}$ . The introduction of the parameter  $\epsilon_F$  also carries over into the energy barriers  $\Delta E_i(\epsilon_F, n)$  and attempt rates  $\nu_{0,i}(\epsilon_F, n)$  that are derived from the assumption of an Arrhenius type behaviour,

$$k_i(\epsilon_F, n, T) = \nu_{0,i}(\epsilon_F, n) e^{-\Delta E_i(\epsilon_F, n)/k_B T}. \quad (\text{IV.23})$$

To arrive at a model that is thermodynamically consistent and allows for interpolation between the  $\epsilon_F$  values that we are probing, we introduce some constraints on the modelling of the energy barriers and attempt rates (similar to the discussion in Sec. II.3.3). For the bond counting approach we assume a linear dependence of the base energy barrier with the parameter  $\epsilon_F$ ,

$$\Delta E_i(\epsilon_F, n) = \Delta E_{i,0} + m_i \epsilon_F + (n - \hat{n}) E_B, \quad (\text{IV.24})$$

where  $E_{i,0}$  is the base energy barrier,  $m_i$  the slope of the linear dependence and  $E_B$  is the effective binding energy, which we have set equal to 235 meV for compatibility of the resulting rate model with the Simple model of Chapter III. For a simple modelling of the attempt rates  $\nu_{0,i}(\epsilon_F, n)$  for edge diffusion and dissociation, we assume that they are constants with respect to  $\epsilon_F$  but can change with transition type  $i$  and with the number of initial neighbours  $n$ ,

$$\nu_{0,i}(\epsilon_F, n) = \nu_{0,i}(n) \quad i \in \{\text{ED-A, ED-B, Diss}\}. \quad (\text{IV.25})$$



As the edge diffusion transitions at A and B step edges, ED-A and ED-B, are going to be the only transitions for which we gather data from two different initial states ( $n = 1$  and  $n = 2$ ), we require the detailed balance condition,

$$\frac{k_{\text{ED-A}}(\epsilon_F, 2, T)}{k_{\text{ED-A}}(\epsilon_F, 1, T)} = \frac{k_{\text{ED-B}}(\epsilon_F, 2, T)}{k_{\text{ED-B}}(\epsilon_F, 1, T)} \quad (\text{IV.26})$$

to ensure thermodynamic consistency. By Inserting Eqs. (IV.23) and (IV.24) into (IV.26), we obtain

$$\frac{\nu_{0,\text{ED-A}}(2)}{\nu_{0,\text{ED-A}}(1)} = \frac{\nu_{0,\text{ED-B}}(2)}{\nu_{0,\text{ED-B}}(1)} := c_{12}. \quad (\text{IV.27})$$

as a condition for the corresponding attempt rates. Only in the case of the ascension transition are we using a linear model function

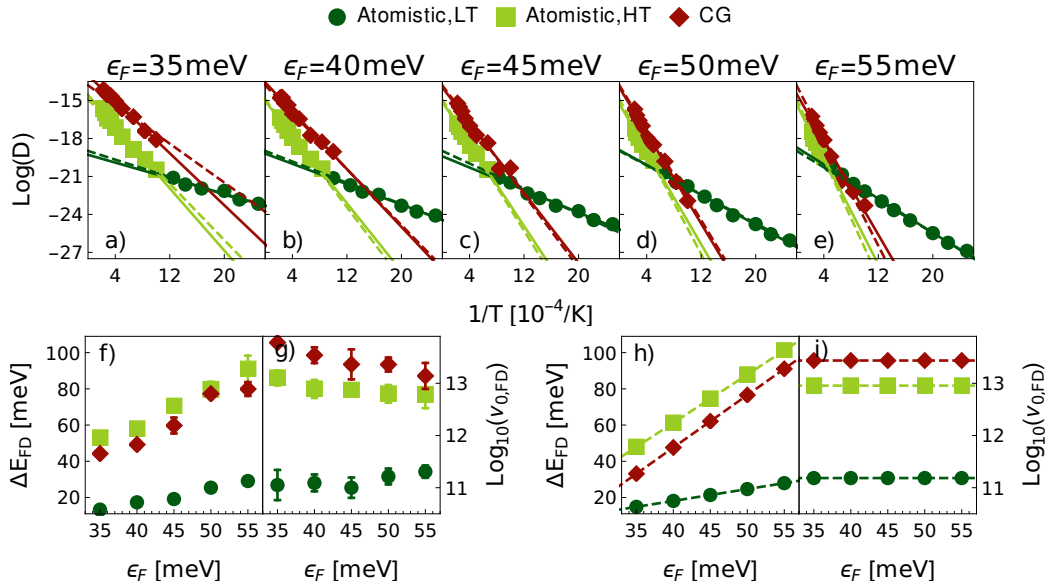
$$\nu_{0,\text{Asc}}(\epsilon_F, n) = y + m_{0,\text{Asc}} \cdot \epsilon_F \quad (\text{IV.28})$$

to give more flexibility to the ascension barriers and allow them to be equal to separately measured potential energies  $E_{\text{Pot}}$  (a justification will be given in the corresponding section).

While we could in principle model the free diffusion process from the last section in conjunction with the dissociation process (as we did in Chapter III), we are not going to do that here because of the discrepancy with the experiments that we have seen in the last section. Instead, we are going to model the free diffusion process separately here as a first example. The Arrhenius plots for the translational diffusion coefficients of both the atomistic and coarse-grained model are shown in Fig. 56(a-e) along with the resulting constrained model [Fig. 56(h,i)] and the unconstrained Arrhenius fit parameters [Fig. 56(f,g)] for comparison. We can nicely see how the constrained model (dashed lines in 56(a-e)) produces a similar behaviour as the unconstrained Arrhenius fits (solid lines) with only a few exceptions like the behaviour of the coarse-grained model at  $\epsilon_F = 35$  meV. The resulting rate model parameters of the constrained model

**Table 9:** Rate model parameters for the free diffusion process derived from the translational diffusion coefficients. Listed are the parameters of Eq. (IV.24) together with the attempt frequencies (calculated with the assumption of single jumps of length  $l = 0.386$  nm).

Model	i	$\Delta E_{i,0}$ [meV]	$m_i$	$\nu_{0,i}$ [GHz]
Atom.	FD-LT	-8.07	0.6546	153
	FD-HT	-45.33	2.664	9025
CG	FD	-68.23	2.896	27301



**Figure 56:** Comparison of unconstrained modelling (via Arrhenius fits) with the constrained modelling for the free diffusion process. (a-e) Arrhenius plots of the translational diffusion coefficients of the coarse-grained and atomistic model. The solid lines are unconstrained Arrhenius fits, while the dashed lines are from the constrained model. (f,g) Energy barriers and attempt rates obtained from the unconstrained Arrhenius fits. (h,i) Constrained model where the functional dependencies are confined to linear/constant via Eqs. IV.24 and IV.25.

are listed in Tab. 9. When using the listed attempt rates in a KMC simulation, it is important to note that they were calculated with the assumption of single jumps between neighbouring sites of length  $l = 0.386 \text{ nm}$  and have to be adjusted according to the implemented geometry. With the linear models of the low and high-temperature regimes of the atomistic model, we are now also able to express the crossover temperature  $T_{\text{Cross}}$  in terms of the interaction strength  $\epsilon_F$ ,

$$T_{\text{Cross}}(\epsilon_F) = -106.027 \text{ K} + 5.7187 \text{ KmeV}^{-1} \cdot \epsilon_F. \quad (\text{IV.29})$$

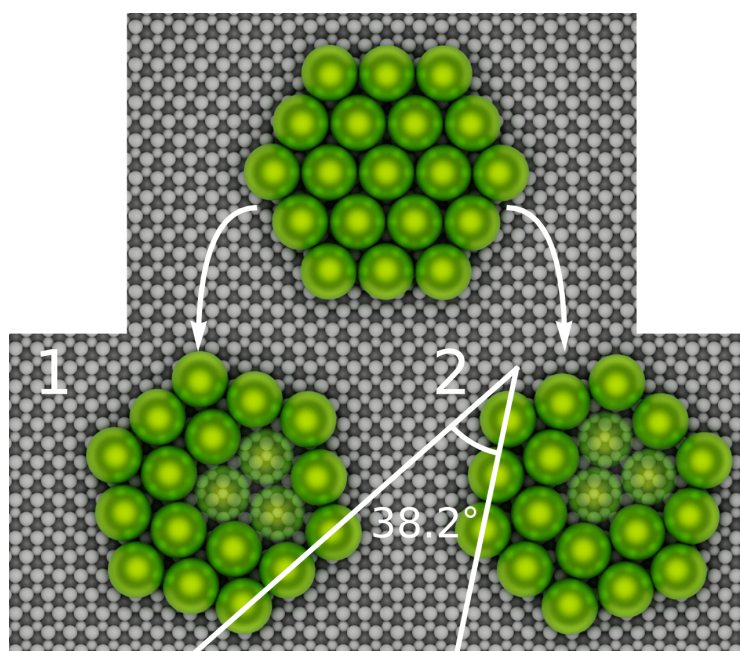
For the discussion of the individual transition types in the following section, we are going to provide similar comparisons of constrained and unconstrained models as in Fig. 56, but the resulting modelling parameters are going to be compiled in the summarizing section IV.1.5.

#### IV.1.4 EDGE DIFFUSION

After the extensive discussion of the dynamics of a single  $C_{60}$  molecule on an otherwise empty  $CaF_2(111)$  surface, we now turn to transitions of single molecules that are bound to clusters of  $C_{60}$  molecules on the substrate. To this end, we increase the size of the simulation box to  $80 \text{ \AA} \times 92 \text{ \AA}$  to make room for a 19-membered hexagonal core cluster to which we can attach additional adparticles at the corners or edges.

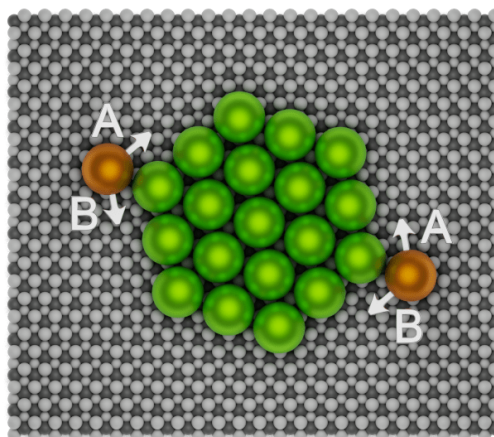
Experiments have shown that  $C_{60}$  clusters on  $CaF_2(111)$  can have two distinct orientations with respect to the substrate [36]. To determine how we can initialize the clusters on the substrate in a stable configuration, we run a few MD simulations with the core cluster aligned parallel to the substrate [Fig. 57 top] at  $T = 300 \text{ K}$  and observe into which alignments it relaxes. As proposed by Ref. [36], we found the same two possible alignments with an angle of  $38.2^\circ$  between them, as shown in Fig. 57 (orientations 1 and 2). The transparent molecules of Fig 57 also expose how the individual adparticles of the cluster all take positions above the 3F lattice sites, which we had also shown to be the preferable position of single molecules on the substrate in Sec. IV.1.2. As the two alignments are basically mirror images of each other, they are equivalent and we choose to initialize the clusters of all following simulations with orientation 1.

**Figure 57:** Two observed cluster orientations after relaxation of a cluster that is initially aligned parallel to the substrate orientation. Three molecules of the final configurations are rendered transparently to expose their position above the 3F lattice site.



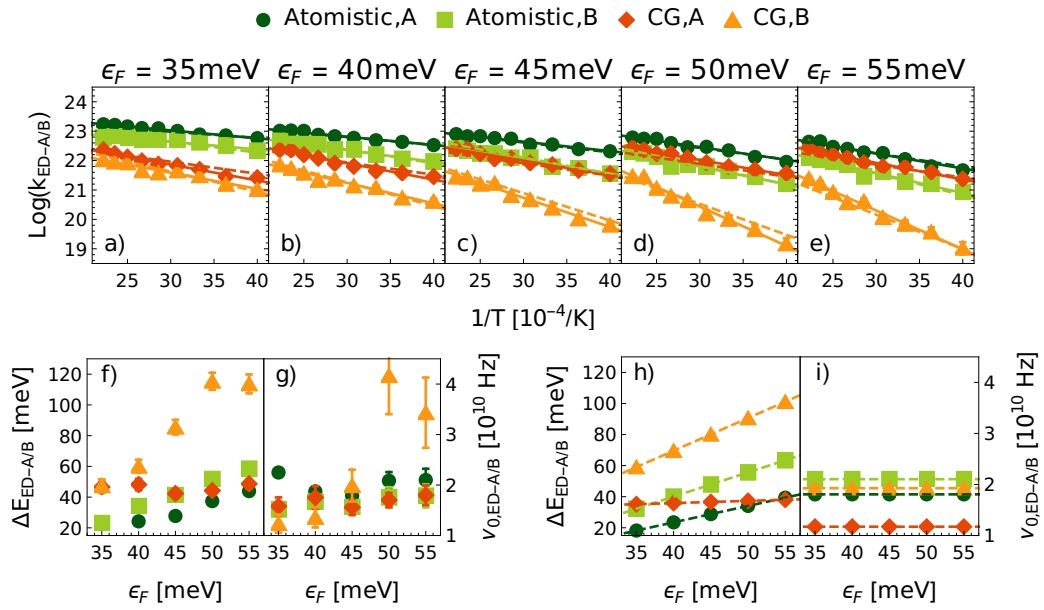
IV.1.4.1 *One Initial Neighbour*

For the observation of edge diffusion with one initial neighbour, we position two  $C_{60}$  molecules on opposite corners of the core cluster such that they are only bound to a single cluster molecule and are on a stable lattice site of the substrate [Fig. 58]. The initialization of the simulations is again conducted via the application of a Langevin thermostat and velocity rescaling to achieve random initial velocities at given temperatures in the range of  $T \in [225, 450]$  K. The coordination number of the two corner molecules is tracked to be able to notice when a transition occurred (as it changes from one to zero or two). After a transition is triggered, a few snapshots are taken in 1 ps intervals to be able to categorize the transitions later and the simulation time is recorded before starting a new trajectory with different initial velocities. One of the goals of these edge diffusion simulations is to test if there is a preferred direction for edge diffusion (as is the case for A and B step edge diffusion for  $C_{60}$  on  $C_{60}(111)$ ), so we categorize the trajectories into transitions towards A and B step edges as shown in Fig. 58. While dissociation from the cluster is in principle also a possible transition in this configuration, the edge diffusion transitions dominate in this medium temperature regime, which is why we are not able to sufficiently sample the dissociation process.



**Figure 58:** Initial configuration for edge diffusion with one initial neighbour. Particles of interest are highlighted in orange.

The Arrhenius plots for this set of simulations are shown in Fig. 59(a-e), showing an overall higher transition rate in the atomistic model and a preference of both models to diffuse into the direction of the A step edge. The difference between the transition rates into the A and B directions increases with  $\epsilon_F$  and becomes especially pronounced in the coarse-grained model. Again, the constrained model (dashed lines in Fig. 59(a-e)) reproduces very well the original Arrhenius data and is close to the unconstrained Arrhenius fits (solid lines) without any significant deviations. In comparison to the free diffusion model, we can note that the energy barriers fall into similar ranges, with the ED-A barrier of the atomistic model being very similar to the free diffusion barrier at low temperatures. Also the free diffusion attempt rate of the atomistic model at low temperatures (when adjusted for the longer jump length of the edge diffusion moves,  $153 \text{ GHz} \cdot (0.386/1)^2 = 2.3 \times 10^{10} \text{ Hz}$ ) is comparable with the attempt

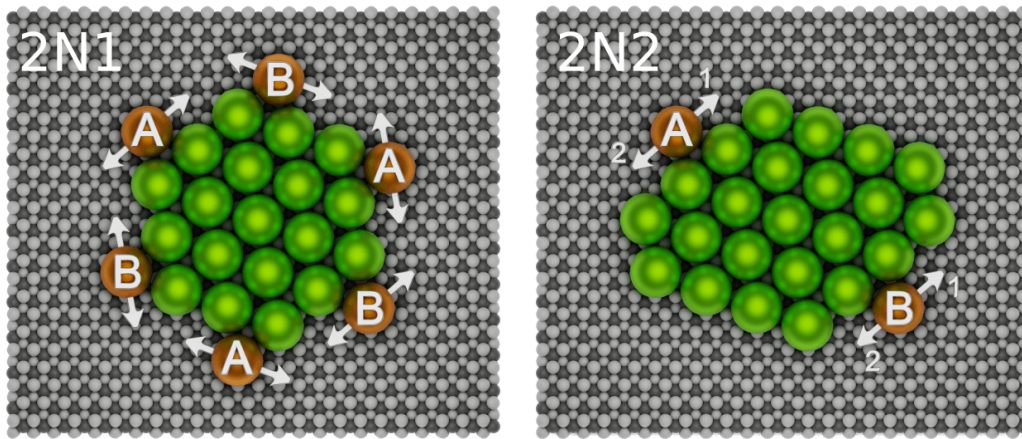


**Figure 59:** Comparison of unconstrained modelling (via Arrhenius fits) with the constrained modelling for the edge diffusion processes with one initial neighbour ( $n = 1$ ). (a-e) Arrhenius plots of the transition rates for the coarse-grained and atomistic model. The solid lines are unconstrained Arrhenius fits, while the dashed lines are from the constrained model. (f,g) Energy barriers and attempt rates obtained from the unconstrained Arrhenius fits. (h,i) Constrained model where the functional dependencies are confined to linear/constant via Eqs. IV.24 and IV.25.

rates obtained here. Since our modelling established a connection between the edge diffusion transitions with one and two initial neighbours via Eqs. IV.24 and IV.27, the constrained model presented in Fig. 59(h,i) also incorporates data from the following set of two neighbour simulations.

#### IV.1.4.2 Two Initial Neighbours

For the case of transitions with two initial neighbours, we set up two distinct configurations. In the first configuration [2N1 of Fig. 60], we place six  $\text{C}_{60}$  molecules at all edges of the core cluster in positions where all relevant transitions can be observed (ascension, dissociation and edge diffusion with one or two final neighbours). The second configuration with a slightly larger core cluster [2N2 in Fig. 60] serves the main purpose as a test to check if there is an asymmetry of the edge diffusion transition with regards to the direction of the transitions (because of the asymmetry given by the angled cluster-substrate alignment). As no asymmetry is observed in this configuration (the transition rates in directions  $A_1$  and  $A_2$ , as well as  $B_1$  and  $B_2$  of Fig. 60 2N2 are the same), the results of both configurations of Fig. 60 are merged for the following discussion.

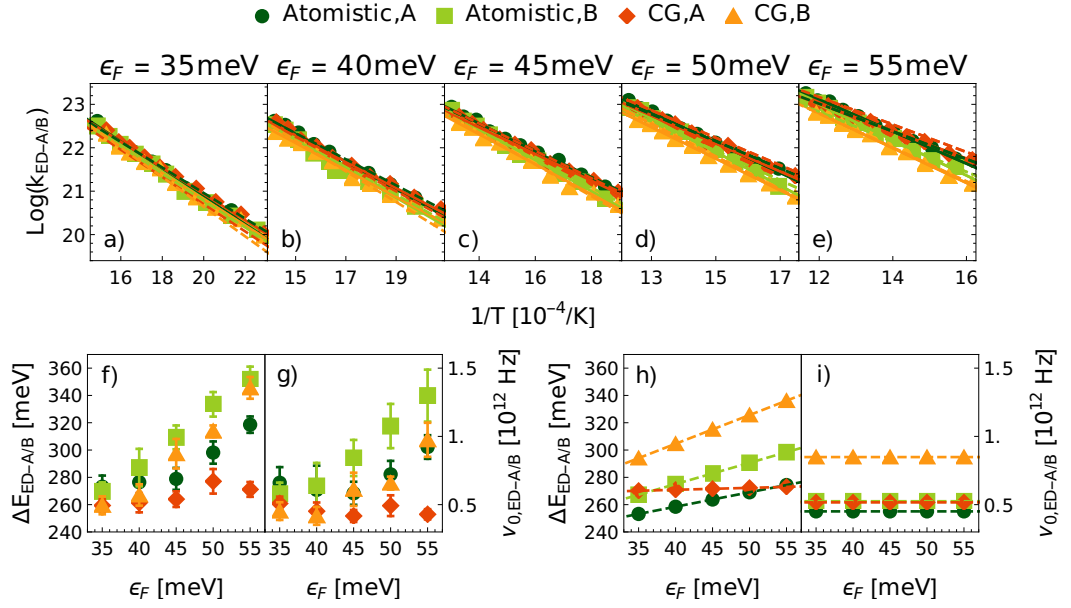


**Figure 60:** Initial configurations for edge diffusion with two initial neighbours. Particles that can go through transitions of interest are highlighted in orange.

Since the coordination number between the  $C_{60}$  molecules in the cluster does not necessarily change after a transition in this case (edge diffusion can go from a two neighbour to another two neighbour state), we have to adjust the transition detection mechanism. To be able to notice transitions we place additional non-interacting “dummy” particles at the initial locations of the adparticles and track the coordination number between those dummy particles and the molecules of interest.

To also gather enough statistics on the rarer dissociation and ascension transitions, we exhaust the probed temperature range to a maximum at which the initial configuration can stably be equilibrated. Since increasing  $\epsilon_F$  also increases the stability of the core cluster, we vary the temperature range in this set of simulations starting with  $T \in [450, 720]$ K at  $\epsilon_F = 35$  meV and increasing it up to  $T \in [650, 920]$ K at  $\epsilon_F = 55$  meV.

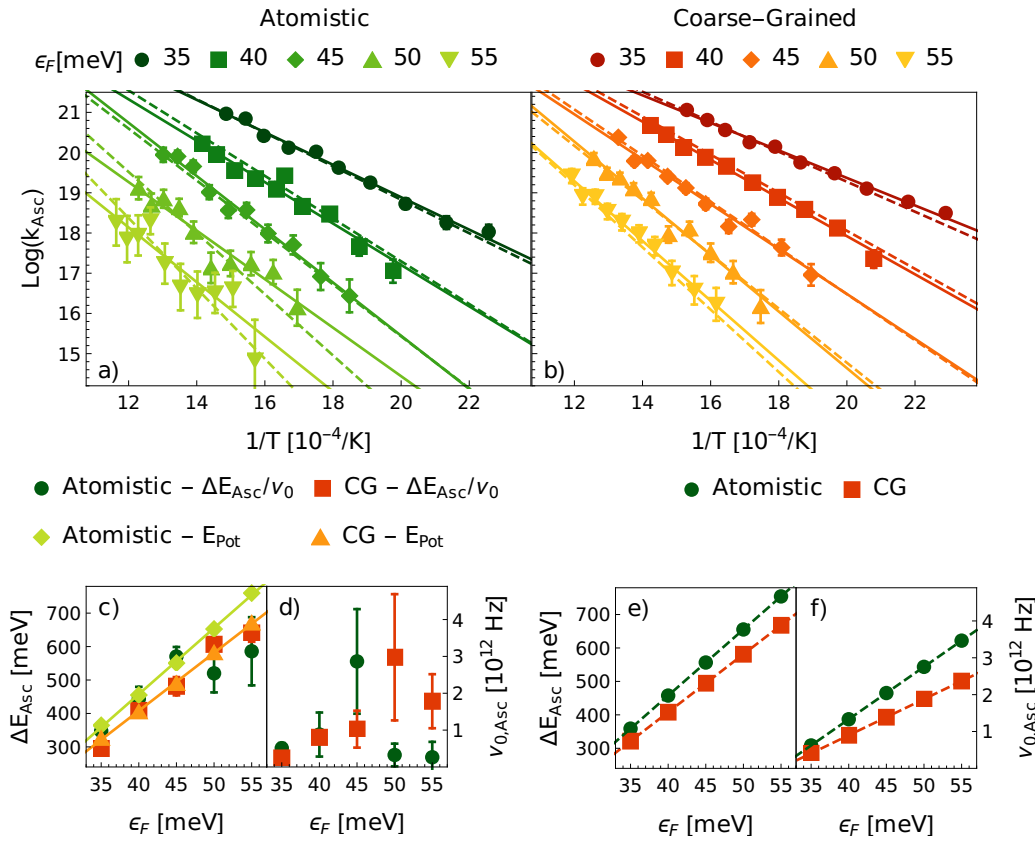
Looking at the Arrhenius plots for the edge diffusion transitions in Fig. 61, we can already see that there is some difference to what we have seen in the one-neighbour edge diffusion case (Fig. 59). In this configuration, the two models now seem to show a much more similar behaviour, starting with almost equal transition rates for A and B step edge diffusion at the low interaction strengths ( $\epsilon_F = 35$  meV). Increasing  $\epsilon_F$  increases the divide between A and B step transition rates very similarly in both models, however at  $\epsilon_F = 55$  meV a stronger divide can clearly be seen in the coarse-grained model. The observation of this agreement of the two models again supports the idea that at higher temperatures the increased rotational diffusion of the atomistic model leads to an effective behaviour similar to the one generated by the rotationally averaged interaction of the coarse-grained model (as discussed in Sec. IV.1.2). However, the agreement could also be explained by the increased coordination of the probed particles, which may diminish the atomistic model’s ability



**Figure 61:** Comparison of unconstrained modelling (via Arrhenius fits) with the constrained modelling for the edge diffusion processes with two initial neighbours ( $n = 2$ ). (a-e) Arrhenius plots of the transition rates for the coarse-grained and atomistic model. The solid lines are unconstrained Arrhenius fits, while the dashed lines are from the constrained model. (f,g) Energy barriers and attempt rates obtained from the unconstrained Arrhenius fits. (h,i) Constrained model where the functional dependencies are confined to linear/constant via Eqs. IV.24 and IV.25.

to lower its effective energy barriers during transitions through adjustment of its orientation (with too many forces pulling on its atoms from different directions). While the difference between the unconstrained [Fig. 61(f,g)] and the constrained modelling [Fig. 61(h,i)] seems more pronounced in this case than for the one-neighbour edge diffusion, the resulting predicted transition rates [dashed lines in Figs. 61(a-e)] are still close to the measured data.

Turning to the ascension transitions with two initial neighbours, the corresponding Arrhenius plots are shown in Fig. 62(a,b). In an ascension transition, a cluster particle usually maintains the bonds to its lateral neighbours while breaking its bond to the substrate. This image of the transition naturally suggests that the energy barrier for ascension strongly increases with the interaction strength  $\epsilon_F$  and should have a value close to the total potential energy of a molecule sitting on the substrate. This is reflected in the Arrhenius plots as we can see how the transition rate and the energy barriers (slope of the fits) for ascension are strongly affected by the value of  $\epsilon_F$ . Additionally, we observe a slightly higher transition rate (and lower energy barrier) in the coarse-grained model, which can be explained by the weaker adparticle-substrate binding energy of the coarse-grained model (See.  $E_{\text{min}}$  values of Tab. 8, Sec. IV.1.2.1). To test how well the ascension barriers  $\Delta E_{\text{ASC}}$  obtained from the Arrhenius fits align

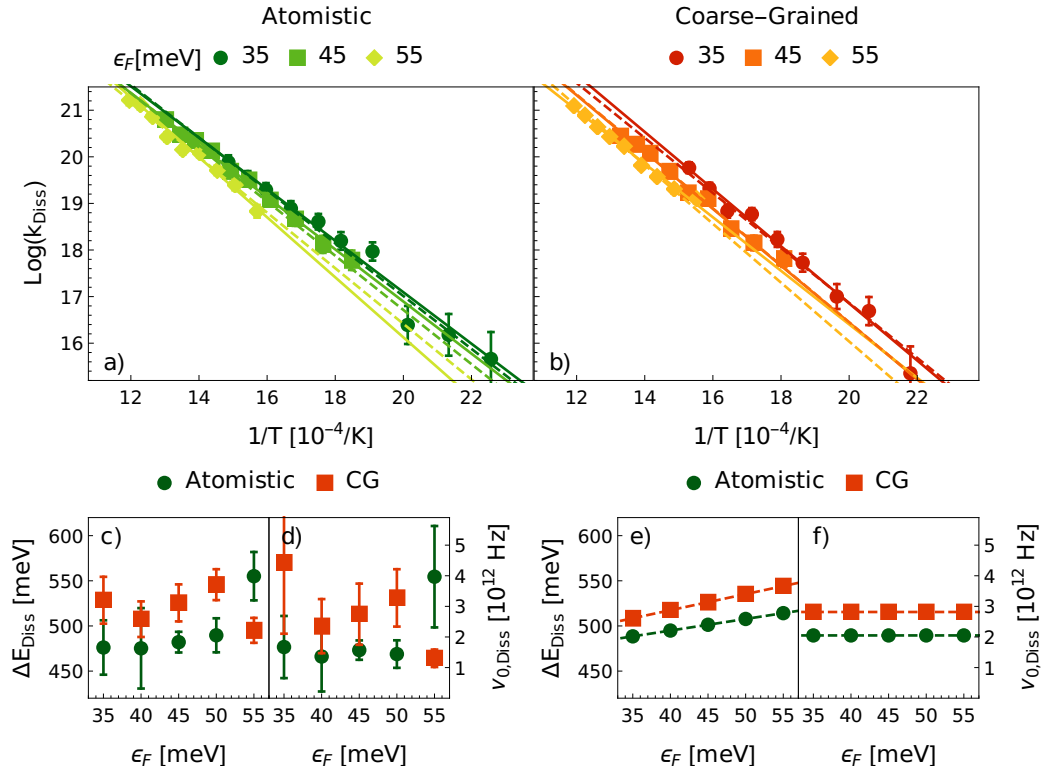


**Figure 62:** Comparison of unconstrained modelling (via Arrhenius fits) with the constrained modelling for the ascension process with two initial neighbours ( $n = 2$ ). (a,b) Arrhenius plots of the transition rates for the coarse-grained and atomistic model. The solid lines are unconstrained Arrhenius fits, while the dashed lines are from the constrained model. (c,d) Energy barriers and attempt rates obtained from the unconstrained Arrhenius fits. Also plotted are potential energies  $E_{\text{Pot}}$  measured for a single  $C_{60}$  molecule on the  $CaF_2(111)$  surface. (e,f) Constrained model where the energy barriers are given by the potential energies  $E_{\text{Pot}}$  of graph (c) and the functional dependence of the attempt rates is confined to linear via Eq. IV.28.

with the actual binding energies/potential energies  $E_{\text{Pot}}$  of  $C_{60}$  on  $CaF_2(111)$ , we have gathered additional data from a small set of simulations where we have measured the mean potential energy of a single  $C_{60}$  molecule on the substrate (In a configuration as shown in Fig. 47) at a temperature of  $T = 300$  K. The results are plotted alongside the Arrhenius parameters in Fig. 62c where we can see that they almost perfectly align with each other except for two strong deviations of the atomistic model at  $\epsilon_F = 50/55$  meV. However, at these two values of  $\epsilon_F$  the discrepancy can be explained by insufficient statistics (e.g. with only 39 observations of ascension at  $\epsilon_F = 55$  meV over the whole temperature range).

We take this finding of the overlap between the  $\Delta E_{\text{Asc}}$  and  $E_{\text{Pot}}$  values as a reason to slightly change our modelling approach for the Ascension transition.





**Figure 63:** Comparison of unconstrained modelling (via Arrhenius fits) with the constrained modelling for the dissociation process with two initial neighbours ( $n = 2$ ). (a,b) Arrhenius plots of the transition rates for the coarse-grained and atomistic model. The solid lines are unconstrained Arrhenius fits, while the dashed lines are from the constrained model. (c,d) Energy barriers and attempt rates obtained from the unconstrained Arrhenius fits. (e,f) Constrained model where the functional dependencies are confined to linear/constant via Eqs. IV.24 and IV.25.

Instead of fitting the energy barriers as linear functions to the Arrhenius data, we use linear fits to the measured potential energies  $E_{\text{pot}}$  (which we can measure with very high precision) and therefore give the freedom of a linear dependence to the attempt rates via Eq. IV.28 to be able to nicely fit the Arrhenius data. The resulting constrained model [Fig. 62(e,f)] very nicely fits to all the Arrhenius data sets [dashed lines in Fig. 62(a,b)] including the atomistic data at  $\epsilon_F = 50/55$  meV, even though it deviates strongly from the unconstrained Arrhenius fits at those values.

Finally, the resulting Arrhenius plots for the dissociation transition with two initial neighbours are shown in Fig. 63. For dissociation, a particle has to mostly overcome the two bond energies to its lateral neighbours (which are independent of  $\epsilon_F$ ), so it can be expected that the dissociation rate only weakly depends on  $\epsilon_F$ . This expectation is reflected in the data shown in Fig. 63. Combining this observation with the relatively low amount of statistics on the dissociation rate, the resulting unconstrained Arrhenius parameters [Fig. 63(c,d)] are all

over the place and by themselves would not suggest any specific functional dependence. In contrast, the constrained model nicely produces the slightly increasing energy barriers of  $E_{\text{DisS}}$  with  $\epsilon_F$  that we would expect (similar in slope as the increase of the free diffusion and edge diffusion barriers).

#### IV.1.5 SUMMARY

In the previous sections, we have conducted and analyzed in detail molecular dynamics simulations on various diffusive transitions of  $C_{60}$  on the  $CaF_2(111)$  surface with two distinct representations of the  $C_{60}$  molecule. We have compared an atomistic representation of the  $C_{60}$  molecule with Lennard-Jones C-F and Buckingham C-Ca interactions to a coarse-grained model of  $C_{60}$  for which we have derived central-body  $C_{60}$ -F and  $C_{60}$ -Ca interactions analogous to the Girifalco potential. The force-field parameters for the implemented interaction potentials were taken from the literature [151] except for the carbon-fluoride interaction energy  $\epsilon_F$  which was left as a single free parameter of the models to be able to tune the molecule-substrate interaction strength.

In the detailed analysis of the free diffusion process in Sec. IV.1.2 we have found that the two models can produce significantly different results. Especially the mean-squared displacement analysis of the free diffusion trajectories at low temperatures and the calculated minimum energy paths make the differences very apparent. Going to high temperatures, the mean-squared displacement analysis and the measurements on the edge diffusion process at high temperatures suggest that the difference between the two models diminishes. These observations can be attributed to the atomistic model's rotational degrees of freedom, which allow it to optimize its alignment to other particles during diffusive transitions to lower the effective energy barriers of the transitions. This ability is nullified at higher temperatures, where the rotational diffusion of the atomistic molecule becomes too fast for any fine alignments to occur, and the effective atomistic interaction more closely represents the rotationally averaged coarse-grained interaction. We were able to determine specific crossover temperatures in the rotational diffusion coefficient at  $T_{\text{Cross,R}} = 163(1)$  K and in the translational diffusion coefficients at  $T_{\text{Cross,T}} \in [100, 200]$  K (depending on the value of  $\epsilon_F$ , see Eq. IV.29), where the atomistic model significantly changes its temperature dependence and starts to approach the behaviour of the coarse-grained model [Fig. 56(a-e)].

While the free diffusion barriers obtained in Sec. IV.1.2 ( $\Delta E_{\text{FD}} \in [15, 90]$  meV) were found to be at variance with the experimental value of 214(16) meV [36] (calculated indirectly from cluster densities of MBE experiments), we determined that the discrepancy can be explained by an effective diffusion process that is hindered by impurity trapping [85] (air/water molecules adsorbed on the

Model	i	n	$\Delta E_{i,0}[\text{meV}]$	$m_i$	$\nu_{0,i}(n)[\text{GHz}]$
Atom.	FD-LT	0	-8.07	0.6546	153
	FD-HT	0	-45.33	2.664	9025
	ED-A	1	-18.9	1.06	18.0
	ED-A	2			450
	ED-B	1	-22.3	1.56	21.0
	ED-B	2			525
	Diss	2	-26.5	1.29	2050
	Asc	2	-333	19.8	$-4355 + 142 \frac{\text{GHz}}{\text{meV}} \epsilon_F$
CG	FD	0	-68.23	2.896	27301
	ED-A	1	30.3	0.142	11.8
	ED-A	2			518
	ED-B	1	-14.8	2.13	19.6
	ED-B	2			861
	Diss	2	-23.8	1.78	2816
	Asc	2	-285	17.3	$-3025 + 98 \frac{\text{GHz}}{\text{meV}} \epsilon_F$

**Table 10:** Final model parameters derived from our MD simulations. Listed are the parameters of Eq. (IV.24) together with the attempt frequencies.

surface acting as nucleation sites). We, therefore, went on to systematically probe edge diffusion/dissociation/ascension transition rates to be able to build comprehensive rate models for the atomistic and coarse-grained representation that will be tested against experimentally observed cluster morphologies in KMC simulations in the next section.

The resulting model parameters are summarized in Tab. 10. The ratios between the attempt rates with one and two initial neighbours (Eq. II.29) are

$$c_{12,\text{Atom.}} = 25.0 \quad c_{12,\text{CG}} = 43.9 \quad (\text{IV.30})$$

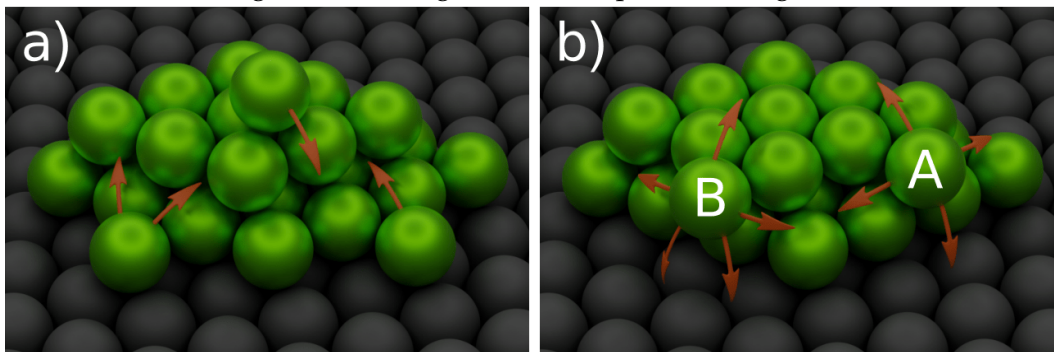
We have demonstrated that our approach of building rate models for KMC simulations based on MD data can also be applied to a system on which no established interaction potentials are available by using a variable molecule-substrate interaction potential with a free parameter (in our case  $\epsilon_F$ ) that is passed on to the resulting rate models to be used as a free parameter of the KMC simulation. While the  $C_{60}$  on  $\text{CaF}_2(111)$  system was computationally much more expensive to simulate than the  $C_{60}$  on  $C_{60}(111)$  system (mainly because of the fine-grained substrate with long-ranged electrostatic interactions), we were still able to capture transitions with up to two initial neighbours, covering the most essential elementary transitions.

## IV.2 KMC SIMULATIONS

In this section, we combine the obtained rate model for  $C_{60}$  on  $CaF_2(111)$  diffusion of the previous section with the Simple model of chapter III to test if these models can reproduce experimental observations. While the Simple model even covers very rare transitions, the modelling for the  $CaF_2$  substrate has to be extrapolated to cover transition rates for the rarer transitions (e.g. with  $n > 2$  initial neighbours) using our modelling with energy barriers from Eq. IV.24 and attempt rates via Eq. II.29 (of Sec. II.3.3) (assuming by default  $c_{23} = c_{34} = c_{45} = c_{56} = 1$ ).

As our MD simulation data does not cover transitions to and from overhang sites (sites that are supported by less than three molecules of the layer below, see Fig. 64), we have to estimate the corresponding transition rates based on the measurements we have done. We assume that the ascension and descension rates to overhang sites are the same as the normal ascension and descension rates. Particles on overhang sites can either do edge diffusion moves according to the transition rates we have observed on fully supported  $C_{60}(111)$  sites, or ascend/descend. The ascension/descension from overhang sites is modelled like a free diffusion process with equal transition rates for ascension and descension. Particles on overhang sites in a B step configuration [Fig. 64b] diffuse just like on a  $CaF_2(111)$  surface and therefore are assigned transition rates based on the free diffusion parameters of the Simple model,  $\Delta E_{Asc/Desc, Overhang-B} = \Delta E_{FD} = 192 \text{ meV}$  and  $\nu_0 = 0.25 \text{ THz}$ . The overhang site in an A step configuration is more stable (with four bonds instead of three) and we, therefore, increase the ascension/descension barrier for those sites to  $\Delta E_{Asc/Desc, Overhang-A} = \Delta E_{FD} + E_B = 427 \text{ meV}$ . As this modelling for the overhang transitions is mostly pulled out of thin air, we could be missing some important intricacies, e.g. a decreased ascension barrier from the substrate to overhang sites in comparison to normal

**Figure 64:** Transitions to and from overhang sites. a) Ascension and descension transitions to overhang sites. b) Edge diffusion, ascension and descension transitions from overhang sites in A (right) and B step (left) configurations.



ascension, the presence of an additional Ehrlich-Schwöbel type barrier for ascension from an overhang site to the top of a cluster, or different edge diffusion rates on overhang sites. A detailed analysis of the overhang transition rates in MD simulations may be an interesting topic for a future study.

A basic KMC simulation for the cluster growth of  $C_{60}$  on  $CaF_2(111)$  was previously implemented by Martin Körner with results published in Ref. [53]. A central part of the rate modelling in this study is the idea of "facilitated dewetting", which is implemented in a bond counting approach that differs from ours,

$$\Delta E_i(n) = \Delta E_{i,0} + (n_1 - n_2)E_B/2, \quad (\text{IV.31})$$

considering both the initial and final number of neighbours,  $n_1$  and  $n_2$ . The parameters for this KMC simulation were mainly taken from the available literature at the time except for an ascension barrier " $\Delta E_{12}$ ", which was tuned to reproduce experimental cluster morphologies at  $\Delta E_{12} = 420 \text{ meV}$ . While the triangular two-layered clusters were reproduced in this study, the complex branched structures obtained in Körner's simulations did not fit the experimental observations very well (discussed in Körner's dissertation [161]). Strong discrepancies are mainly observed in the cluster sizes (due to the low free diffusion rates as a result of Loske's diffusion parameters taken from Ref. [36]) and in the branching of the second layer nuclei (due to missing distinction between A and B step edge diffusion in both layers).

While we can improve upon both of those discrepancies with our more detailed modelling, a major point we want to address in the following sections is the understanding that the complex cluster morphologies are induced by particles on the edge of the second layer via the "facilitated dewetting" effect, which was established in Ref. [53]. The way this effect was implemented (Eq. IV.31) suggests a long-range interaction between  $C_{60}$  molecules that enables the facilitation of the ascension transition, incompatible with the known short range of the  $C_{60}$ - $C_{60}$  interaction. Therefore we want to offer an alternative with our modelling and show that facilitated dewetting is not a necessary process for the evolution of the experimentally observed cluster morphologies.

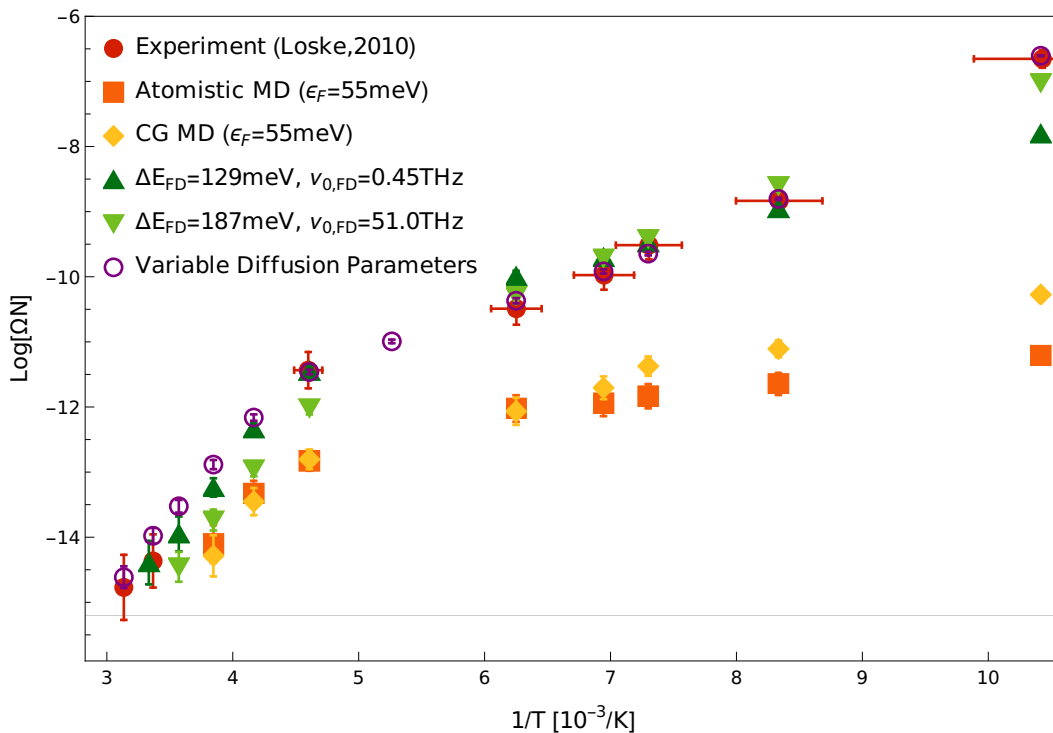
### IV.2.1 CLUSTER DENSITIES

Before we move towards the tuning of the model parameter  $\epsilon_F$  and the comparison of cluster morphologies, we want to make sure that our simulation generates cluster sizes that are comparable to the ones observed in the experiment. Hence, we have to ensure that we reproduce the cluster densities of the experiment. As data for the low-temperature regime, we have the density measurements pub-

lished in Ref. [36]. These were obtained in deposition experiments employing a flux of  $F \approx 0.026$  ML/min over a time of roughly four minutes. To this data set, we add the densities shown in the AFM images at temperatures  $T = 297$  K and  $T = 319$  K published in Ref. [53], which were captured with the same experimental setup after a two minute deposition phase. For a comparison to this experimental data set, we run KMC simulations and measure the resulting cluster densities at temperatures  $T \in \{96, 120, 137, 144, 160, 190, 217\}$  K with a four minute deposition phase and at temperatures  $T \in \{240, 260, 280, 297, 319\}$  K with a two minute deposition phase.

In the discussion on the MD simulations of the free diffusion process (Sec. IV.1.2) we have observed a strong discrepancy between the measured diffusion barriers of the simulations (20 – 90 meV) and the experimental value determined by Felix Loske (214 meV) from cluster density measurements [36]. As we can see in Fig. 65 this discrepancy carries over into the cluster density measurements if we use the diffusion parameters as measured in the MD simulations (orange squares and yellow diamonds). Even using the upper end of our parameter range  $\epsilon_F = 55$  meV, the diffusion process is still way too fast in the whole temperature range, resulting in significantly lower cluster densities compared to the experiment.

**Figure 65:** Cluster density measurements from experiments (red circles) in comparison with measurements from our KMC simulations employing various models for the diffusion parameters. The “Variable Diffusion Parameters” are calculated via Eqs. IV.32 and IV.33.



Trying different combinations of energy barriers and attempt rates for the diffusion process (two example configurations are plotted as dark and light green triangles in Fig. 65), we find that the lower temperature range can be well described with diffusion barriers up to 250 meV (and attempt rates in the order of up to  $10^{17}$  Hz), while the higher temperatures are better reproduced with diffusion barriers approaching 100 meV (and attempt rates in the order  $10^{11}$  Hz). This observation is compatible with our suspicion, that impurities on the surface affect the experimental result. The colder a surface, the stronger it acts as a sink for impurities adsorbed from the imperfect vacuum, resulting in a stronger increase of cluster densities at low temperatures due to the larger amount of nucleation sites. This translates into a larger effective diffusion barrier at low temperatures, which should trend towards the diffusion barrier of a clean substrate at high temperatures.

To be able to best reproduce the cluster densities in the whole temperature range we therefore determined a model (via an initial guess and consecutive adjustments) for the diffusion barrier and attempt rate that linearly depends on the inverse temperature. The result is the following model:

$$\Delta E_{\text{FD}} \left( \frac{1}{T} \right) = 0.061 \text{ eV} + 19.61 \text{ eV K} \cdot \frac{1}{T}, \quad (\text{IV.32})$$

$$\nu_{0,\text{FD}} \left( \frac{1}{T} \right) = \exp \left( 19.26 + 1981 \text{ K} \cdot \frac{1}{T} \right) \text{ Hz}. \quad (\text{IV.33})$$

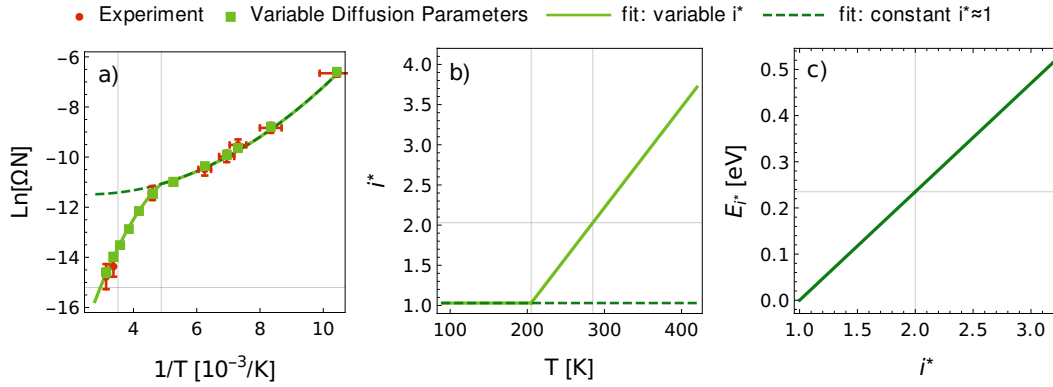
The cluster densities produced by this model are plotted as hollow purple circles in Fig. 65, nicely fitting to the experimental data.

An interesting feature of the experimentally observed cluster densities that is also observed in the simulation is the significant drop between  $T = 217$  K and  $T = 297$  K. This drop is not generated by an increase of the diffusion rate, but by the decreasing stability of small clusters, which hinders nucleation. In nucleation theory, this is modelled via the critical cluster size  $i^*$  and the critical cluster energy  $E_{i^*}$  in the formula (derived in Sec. II.2.3)

$$\Omega N \approx \eta \left( \frac{\Omega^2 F}{D} \right)^{\frac{i^*}{i^*+2}} e^{E_{i^*}/(i^*+2)k_B T} \quad (\text{IV.34})$$

$$D = \frac{1}{4} l^2 \nu_{0,\text{FD}} e^{-\Delta E_{\text{FD}}/k_B T}. \quad (\text{IV.35})$$

At low temperatures, the prediction of Eq. IV.34 fits our simulation results very well for parameters  $\eta \approx 0.13$ ,  $i^* \approx 1.0$  and  $E_{i^*} \approx 0$  meV [green dashed line in Fig. 66a], in line with the expectation of monomers as critical clusters for low temperatures (where dimers are already considered stable). However, starting at temperatures of around 200 K the simulation results start to strongly deviate from this model line. We can recover the behaviour of the simulation by adding



**Figure 66:** Cluster density prediction from nucleation theory with constant and variable modelling of the  $i^*$  parameter.

a temperature dependence to the parameter  $i^*$  and modelling  $E_{i^*}$  via

$$i^*(T) = \begin{cases} \sim 1.0 & T < T_1 \\ \sim 1.0 + \frac{(T-T_1)}{(T_2-T_1)} & T \geq T_1 \end{cases} \quad E_{i^*} = (i^* - 1)E_B, \quad (\text{IV.36})$$

with  $T_1 \approx 205$  K (the temperature at which the deviation from  $i^* \approx 1.0$  starts),  $T_2 \approx 285$  K (the temperature at which  $i^* \approx 2.0$  is reached) and  $E_B = 235$  meV (the binding energy between two  $C_{60}$  molecules). The functional forms of Eqs. IV.36 are plotted in Fig. 66(b,c) and the resulting prediction if Eqs. IV.36 are put into Eq. IV.34 is plotted as a light green line in Fig. 66a

For a final interesting note, from Eqs. IV.32, IV.33 and IV.35 we can derive a transport property  $\tau = l^2/4D$ , which behaves analogously to transport properties of glass-forming liquids [162, 163] where a parabolic form

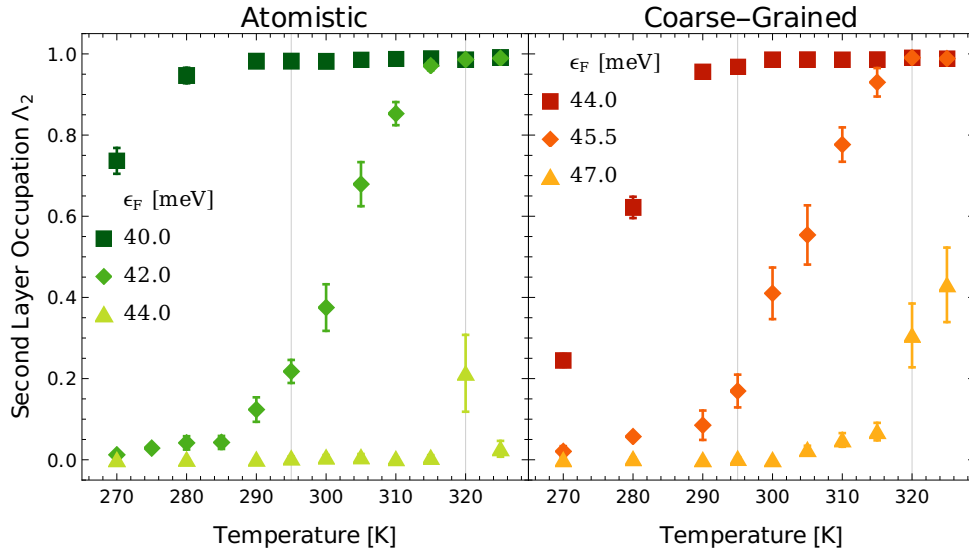
$$\log(\tau/\tau_0) = J^2 \left( \frac{1}{T} - \frac{1}{T_0} \right)^2 \quad T < T_0$$

was observed (also called "super-Arrhenius" behaviour). The corresponding parameters of our modelling are  $J \approx 477$  K,  $T_0 \approx 357.5$  K and  $\log(\tau_0) \approx -21.04$ .

## SECOND LAYER OCCUPATION, $\epsilon_F$ -PARAMETER TUNING

As shown in Felix Loskes experiments at around room temperature [53], the  $C_{60}$  clusters on  $CaF_2(111)$  transition from a mainly one-layered growth with interesting morphologies in the transition regime (observed at  $T = 297$  K) into fully two-layered growth of triangular clusters (observed at  $T = 319$  K). This observation is perfect for the tuning of our one free KMC model parameter,  $\epsilon_F$ ,





**Figure 67:** Second layer occupation after the deposition phase plotted against substrate temperature for the atomistic and coarse-grained model.

as it strongly affects the ascension rate and can therefore only reproduce this experimental observation in a very narrow range of values.

For the parameter tuning, we aim to reproduce the experimental setup with a molecular flux of  $F = 5 \times 10^{-4} \text{ s}^{-1} \text{ nm}^{-2} \approx 0.026 \text{ ML/min}$  and a two minute deposition phase in three-layered simulation boxes of area  $A = 3.464 \mu\text{m}^2$  ( $2000 \times 2000$  unit cells). Simulations are run at temperatures  $T \in [270, 325] \text{ K}$ , varying the interaction parameter in the range  $\epsilon_F \in [40, 50] \text{ meV}$ .

The resulting second layer occupations are shown in Fig. 67 for a selection of  $\epsilon_F$  values. The second layer occupations are calculated immediately after the deposition phase as the fraction  $\Lambda_2 = N_2/N_1$ , with  $N_i$  being the number of particles in the  $i$ -th layer. In both models we observe a high sensitivity of  $\Lambda_2$  to the parameter  $\epsilon_F$ . To reproduce the experimental observation we look for an  $\epsilon_F$  at which full coverage ( $\Lambda_2 \approx 1$ ) of the second layer is reached by  $T = 320 \text{ K}$  and also a low coverage of around  $\Lambda_2 \approx 0.2$  evolves at  $T = 295 \text{ K}$ . For the atomistic model (Fig. 67 left) these criteria are satisfied at  $\epsilon_F = 42.0(5) \text{ meV}$

**Table 11:** Optimal  $\epsilon_F$  values for the two models and resulting energy barriers, attempt rates and transition rates for ascension with two initial neighbours ( $n = 2$ ). For comparison, the values used by Martin Körner [53] in KMC simulations investigating the same system are also listed.

Model	$\epsilon_F [\text{meV}]$	$\Delta E_{\text{Asc}} [\text{meV}]$	$\nu_{\text{Asc}} [\text{THz}]$	$k_{\text{Asc}}(295 \text{ K}) [\text{Hz}]$
Atomistic	42.0(5)	499(10)	1.61(7)	4807
CG	45.5(5)	502(9)	1.43(5)	3795
Körner [53]	-	$\Delta E_{12} = 420$	$\nu_{12} = 1.00$	66791

and for the coarse-grained model (Fig. 67 right) at  $\epsilon_F = 45.5(5)$  meV. The corresponding dewetting energy barriers, attempt rates and overall transition rates (at  $T = 295$  K) are given in Tab. 11.

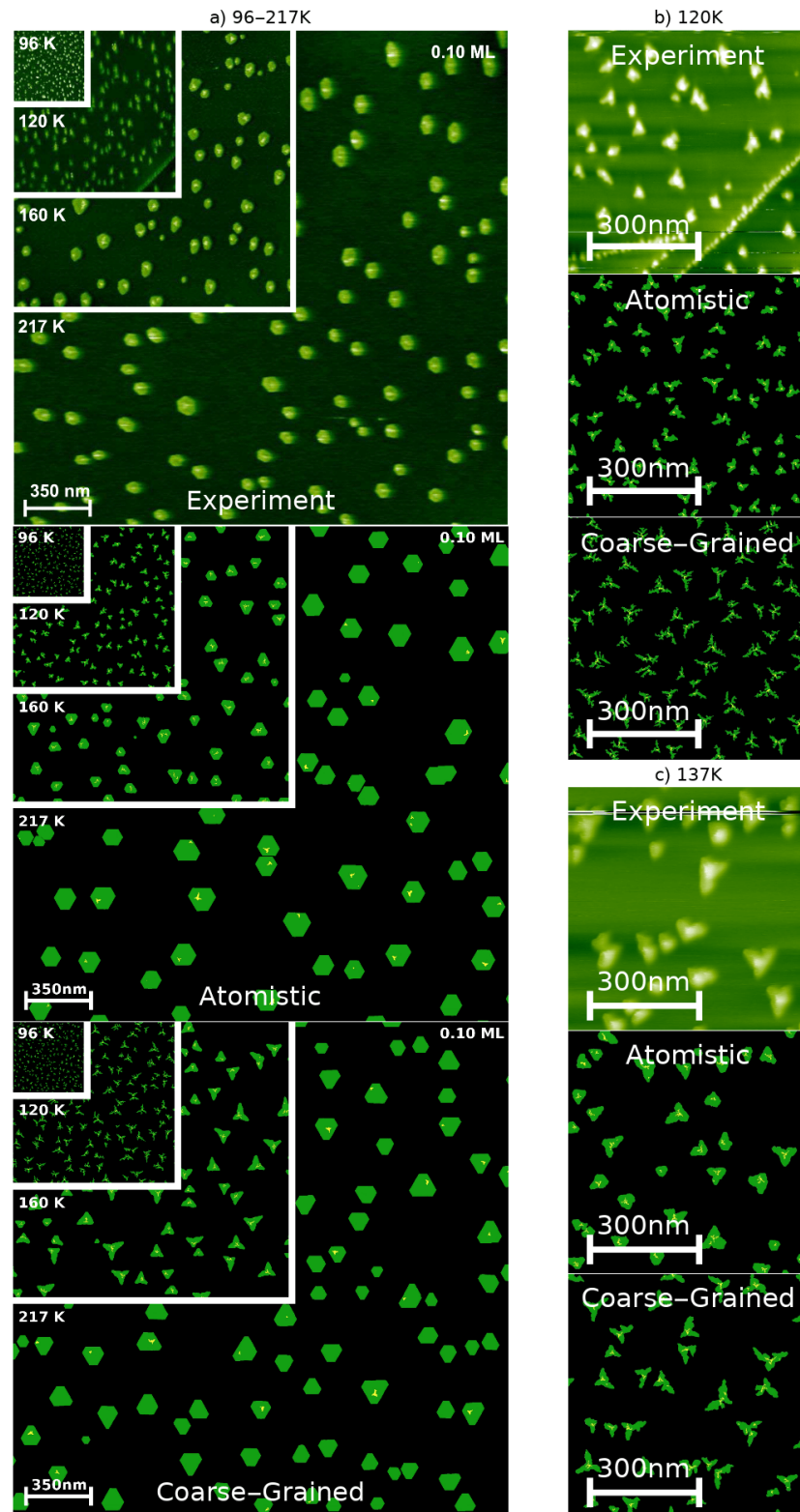
We see that the two different  $\epsilon_F$  values of our two models lead to practically the same energy barrier and attempt rate for the ascension transition (for  $n = 2$ ) of around  $\Delta E_{\text{Asc}} \approx 500$  meV and  $\nu_{\text{Asc}} \approx 1.5$  THz. Comparing those values to the previous work of Martin Körner [53], we see that his assumption of  $\nu_{12} = 1.00$  THz is actually very close to our result, but combined with the lower energy barrier  $\Delta E_{12} = 420$  meV, leads to much higher ascension rates at the temperatures of interest. The main source of this discrepancy may be the choice of parameters for the free diffusion process, which in Martin Körner's model were taken from Ref. [36]. These lead to cluster densities much higher than observed in the experiment and consequently to smaller cluster sizes, needing higher dewetting rates to accomplish comparable morphologies.

## IV.2.2 CLUSTER MORPHOLOGIES

In this section we test our obtained models by comparing the simulation results we get at low and high temperatures in side-by-side comparisons with NC-AFM images from experiments.

For the low-temperature regime (96 – 217 K), we use the NC-AFM image published in Ref. [36] from the set of cluster density measurements. For the lower temperatures of 120 K and 137 K we were kindly provided with additional images from Felix Loske and Angelika Kühnle on which the cluster morphologies can be identified (These images were also taken in 2010 along with the other measurements for the purpose of counting clusters and are therefore not of the best visual quality). For the higher temperatures (297 – 319 K) we take the selection of NC-AFM images published in Ref. [53].

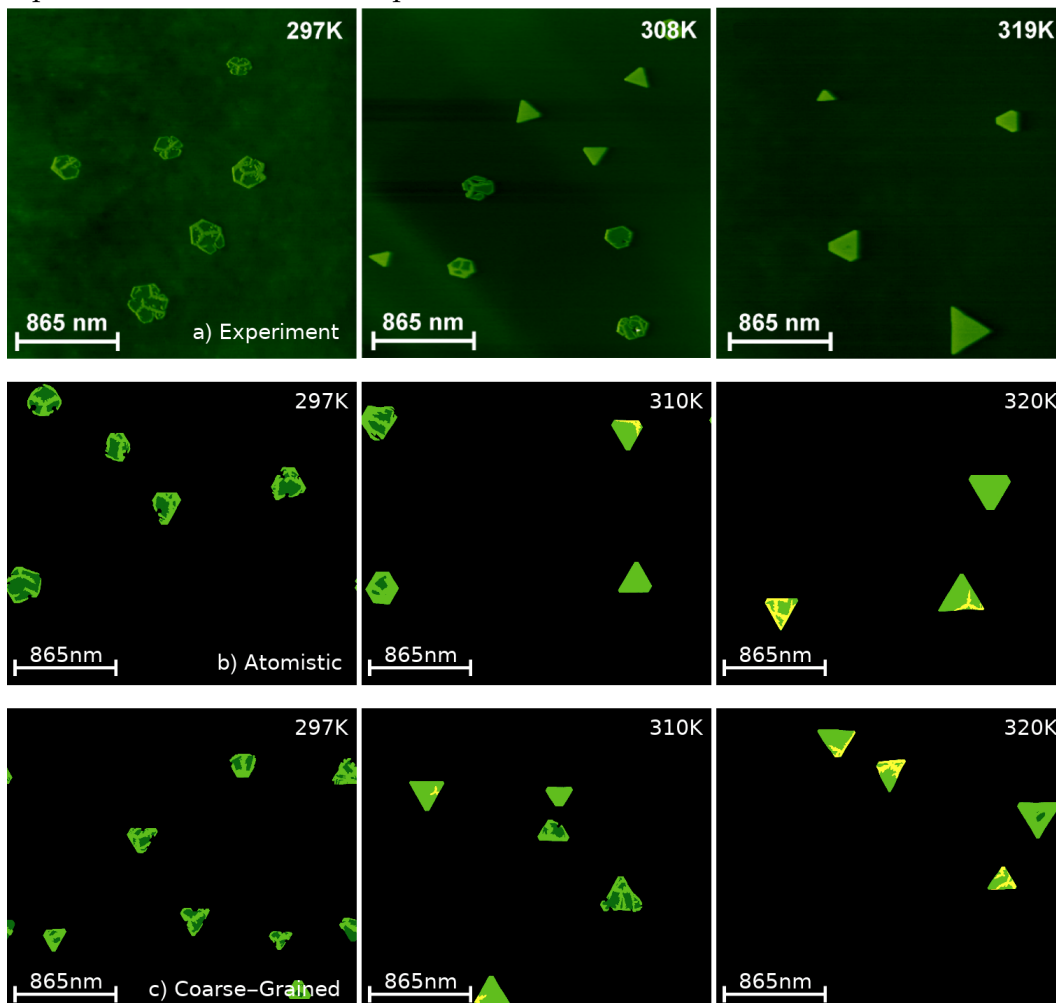
For the simulation snapshots, we aim to reproduce the experimental protocol by applying a molecular flux of  $F \approx 0.026$  ML/min over a four minute deposition phase at substrate temperatures  $T \in \{96, 120, 137, 160, 217\}$  K and a two minute deposition phase at temperatures  $T \in \{297, 310, 320\}$  K. For the lower temperatures (up to 160 K) we run system sizes of  $2000 \times 2000$  unit cells, which we increase to  $3000 \times 3000$  unit cells for the temperatures above 160 K to be able to produce comparable snapshot sizes without periodic images. As rate models for the KMC simulation, we use the atomistic and coarse-grained models with the parameters we have determined in the previous section,  $\epsilon_F = 42.0$  meV (atomistic) and  $\epsilon_F = 45.5$  meV (coarse-grained). To obtain the correct cluster sizes, the free diffusion transition rates are calculated using the parameters given in Eqs. IV.32 and IV.33 for both (atomistic and coarse-grained) models.



**Figure 68:** Comparison of AFM images by Felix Loske from low-temperature deposition experiments with KMC simulation results. Top left experimental images are reprinted with permission from Ref. [36]. Copyright 2010 by the American Physical Society.

The low-temperature results are shown in the side-by-side comparisons of Fig. 68. In subfigure 68a the atomistic model nicely reproduces the compact shapes that have been observed in the experiments at 160 K and 217 K, whereas the coarse-grained model produces triangular star-shaped clusters at 160 K and slightly triangular clusters at 217 K. This suggests that the strong difference between A and B step edge diffusion is exaggerated in the coarse-grained model. On the other hand, the slight difference between the A and B step edge diffusion in the atomistic model seems to produce the correct morphologies in this low-temperature range, which is especially apparent in subfigures 68(b,c). The atomistic model reproduces the slight triangular star shapes at 120 K and 137 K very well, while the coarse-grained model overshoots with pronounced (dendritic) triangular star shapes.

**Figure 69:** Comparison of AFM images by Felix Loske from high-temperature deposition experiments with KMC simulation results. Experimental images are reprinted from Ref. [60] with permission of Felix Loske.



The high-temperature results are shown in Fig. 69 and tell a similar story. The edge diffusion rates of the coarse-grained model lead to the formation of triangular base clusters, which then keep this triangular shape when transitioning into the branched structures with the formation of second layer nuclei [c.f. Fig. 69c 297 K and 310 K]. In contrast, the atomistic model again comes very close to the experimental observation as it forms hexagonal base clusters, which turn into the branched structures with outer layer rims [Fig. 69b 297 K and 310 K]. Also, the transition towards the formation of triangular clusters with fully covered second layers between 310 K and 320 K is reproduced extremely well.

However, there are also minor differences between the results of the atomistic model and the experiments. Firstly, the branched clusters of the experiment [Fig. 69a 297 K] maintain the hexagonal shape of the base clusters to a large degree, while the clusters in the simulation turn into more random shapes. Secondly, the KMC simulation (in both the atomistic and coarse-grained model) prematurely forms third layer nuclei [c.f. Fig. 69(b,c) 320 K], which have not been observed at all in the experiment at this early stage of the deposition process. Both of these differences can potentially be explained by our arbitrary modelling of the transition rates towards and from overhang sites (discussed at the beginning of Sec. IV.2), which is not supported by MD simulations and is most probably wrong in some way.

To summarize this section, we can conclude that the rate model obtained from the atomistic MD simulations, reproduces the experimental observations extraordinarily well in the full temperature range. The evolution of the branched structures with an outer rim was previously explained by the "facilitated dewetting" effect in Ref. [53]. The way this effect was implemented in Ref. [53] is arguably at odds with the short range of the  $C_{60}$ - $C_{60}$  interaction and the results of our modelling (which only considers the initial number of neighbours) show that "facilitated dewetting" is not a necessary process for the evolution of the experimentally observed cluster morphologies.





## PERSPECTIVES

In this thesis, we presented a bottom-up approach for the determination of a thermodynamically consistent transition rate model for KMC simulations based on MD simulation data, applied to the example of the epitaxial growth of  $C_{60}$  on a  $CaF_2(111)$  substrate.

We split the task into two parts by first setting up MD simulations of  $C_{60}$  molecules on a  $C_{60}(111)$  surface, interacting via the coarse-grained Girifalco potential [22, 23]. With this set of coarse-grained simulations, we were able to measure a large portion of possible elementary transitions for the diffusion processes of  $C_{60}$  on  $C_{60}(111)$  (which are assumed to also be valid for the higher layers of  $C_{60}$  cluster growth on other substrates like  $CaF_2(111)$ ). To test the thermodynamic consistency of the derived transition rate models, we have measured the entropy production along the stochastic trajectories in the KMC simulations and we have shown that the naive approach of simply taking the energy barriers and attempt rates from the individual transition rate measurements of the different processes can lead to a thermodynamically inconsistent model with a constant entropy production. We have then derived a thermodynamically consistent constrained "Simple" rate model for  $C_{60}$  on  $C_{60}(111)$  that produces the characteristic triangular star-shaped clusters that have been observed in various  $C_{60}$  epitaxy experiments.

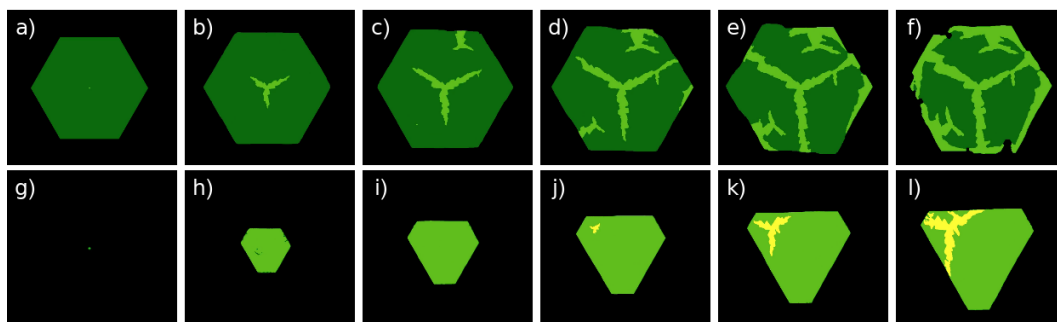
We applied this model to run KMC simulations for multilayer growth of  $C_{60}$  on mica and compared the results to the experimental and simulation data of Ref. [55]. We were able to explain the large discrepancy between the obtained diffusion barrier (for  $C_{60}$  on  $C_{60}(111)$ ) of Ref. [55] and other sources in the literature [52, 54, 138] via the diffusion-hindering effect of grain boundaries that evolve during coalescence of a large number of clusters in the first  $C_{60}$ -layer on the mica substrate. To support this hypothesis we have shown that the evolution of grain boundaries with a high cluster density in the first layer

can significantly increase the cluster density of consecutive layers in our KMC simulations.

For the second part of the rate model, we had to set up MD simulations involving  $C_{60}$  and the  $CaF_2(111)$  substrate, for which we did not have any established interaction potentials. Using the interaction potentials determined in Ref. [151] as reference points, we derived potentials for the necessary C-F and C-Ca interactions in our system with a single free parameter  $\epsilon_F$ . To test the impact of a Girifalco-style coarse-graining of the  $C_{60}$  molecule on the diffusion simulations, we implemented a coarse-grained and an atomistic (rigid body) representation of  $C_{60}$  and ran simulations with both representations to compare the results. For the free diffusion process of  $C_{60}$  on  $CaF_2(111)$ , we observed a crossover in diffusive behaviour of the atomistic model in a temperature range of  $T \in [100, 200]K$ , depending on the choice of the interaction parameter  $\epsilon_F$ . While the coarse-grained and atomistic model exhibited very different temperature scaling at low temperatures (below the crossover temperature of the atomistic model), the difference diminished for higher temperatures. This observation is compatible with the understanding that this type of coarse-graining can be an accurate model for high temperatures, at which the  $C_{60}$  molecule rotates so quickly that the individual carbon positions can be assumed to be smeared out over a spherical surface for any interaction with other atoms or molecules. In comparison to the diffusion barrier determined by Felix Loske via cluster density measurements [36] (214(16) meV), our results (20 – 90 meV) were found to be at variance. However, the probable presence of impurities on the  $CaF_2(111)$  surface in the experiments (especially at low temperatures) can account for this discrepancy and we proceeded to derive transition rate models for both representations of  $C_{60}$  on  $CaF_2(111)$  from transition rate measurements of edge diffusion, dissociation and ascension transitions.

To test the obtained rate models in KMC simulations, we first came up with a modelling for the free diffusion process that reproduces the experimentally observed cluster densities. The free  $\epsilon_F$  parameter of the two models was then tuned to reproduce the transition to the growth of two-layered clusters at  $\sim 320 K$ . The resulting rate models were then used to generate cluster morphologies in the whole experimentally observed substrate temperature range of  $T \in [96, 320]K$ , which were then compared to the observed cluster morphologies of the experiment. While the coarse-grained model produced significantly different morphologies (mostly due to its stronger difference between A and B step edge diffusion), the atomistic model reproduced the experimentally observed cluster morphologies very well across the whole temperature range. We also obtained the interesting two-layered cluster morphologies with the branched structure and outer rims observed at  $\sim 297 K$  and the fully covered triangular clusters at  $\sim 320 K$  without the implementation of a "facilitated dewetting" pro-





**Figure 70:** Growth trajectories at 310K starting with a second layer nucleus at different times of the growth process. a-f) If the second layer nucleus forms at a very late stage of the growth process, the cluster grows into the branched structure with outer rims on the second layer. g-l) With a second layer nucleus at the start of the cluster growth process, it grows into the triangular cluster with a fully covered second layer.

cess, which was previously used to explain the evolution of those morphologies [53].

Summarizing this thesis, we have shown that the epitaxial growth processes of  $C_{60}$  can very well be described by individual elementary transitions in KMC simulations and that rate models for such systems, which can accurately reproduce many details of the experimental system, can be derived from MD simulation data. To build upon our findings, a series of future projects are possible. Firstly, the obtained final transition rate model can be used to further study the exact trajectories that lead to certain morphologies, e.g. one could investigate when and where a second layer nucleus needs to form on top of a base cluster to make it grow into a specific morphology [see Fig. 70 for an example]. A different direction would be to use the KMC simulation to explore the space of possible experimental protocols to try to predict outcomes for experiments that have not been conducted yet. Running KMC simulations (once the model is determined) is arguably much easier and cheaper than doing the actual experiment in a lab and one can potentially expand our ability to manipulate the epitaxial growth process (e.g. via variation of substrate temperature and/or molecular flux) to obtain specific desired results. A direction for a smaller project could be to investigate the transitions to and from overhang sites in MD simulations for the example of  $C_{60}$  on  $CaF_2(111)$  to explore if a more accurate modelling of those transitions (which could involve some light form of "facilitated dewetting") can improve the agreement of the KMC simulations with the experiment even further. Lastly, our approach can also be applied to other systems to derive rate models and investigate them in KMC simulations.



# A

## APPENDIX: KMC IMPLEMENTATION

### CONTENTS

---

A.1	Implemented Geometry and Framework . . . . .	<b>114</b>
A.1.1	Lattice Coarse-Graining . . . . .	<b>114</b>
A.1.2	The GridNode Class . . . . .	<b>116</b>
A.1.3	The RateModel Class . . . . .	<b>117</b>
A.1.4	The Input Class . . . . .	<b>118</b>
A.1.5	The FullereneKMC Class . . . . .	<b>119</b>
A.2	Fast Diffusion . . . . .	<b>120</b>
A.3	Performance . . . . .	<b>126</b>
A.3.1	System size scaling . . . . .	<b>126</b>
A.3.2	Temperature scaling . . . . .	<b>128</b>

---

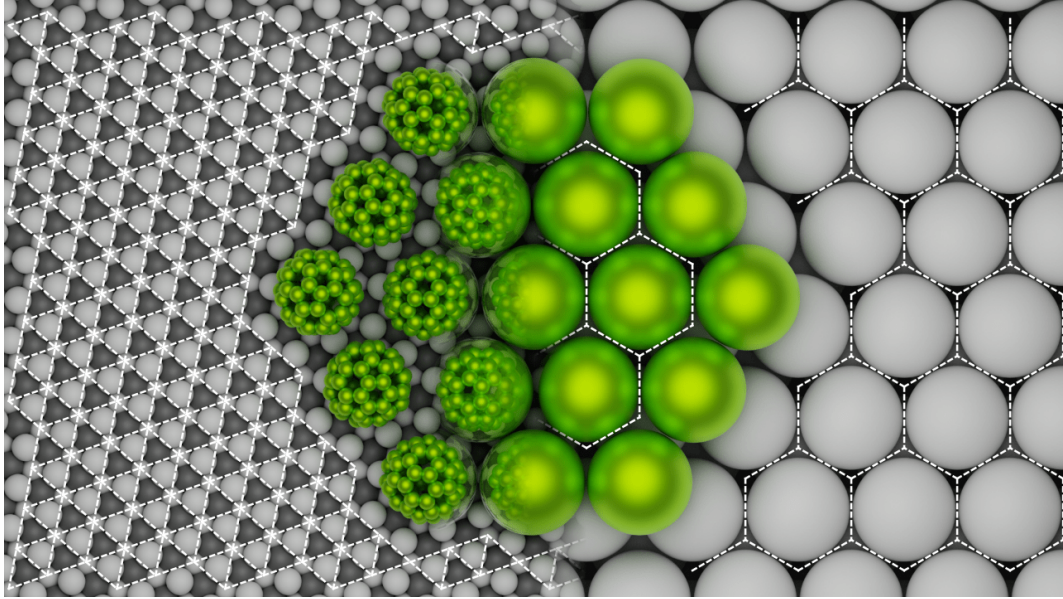
In this chapter, we are going to present some details on the KMC algorithm that we have implemented to simulate the epitaxial growth of  $C_{60}$ . Our code is written in C++ and tested on Linux. It can be accessed via GitLab ([https://gitlab.rlp.net/wjanke/kmc\\_fullerenediffusion](https://gitlab.rlp.net/wjanke/kmc_fullerenediffusion)) or it can be requested via mail ([janke4592@gmx.de](mailto:janke4592@gmx.de)) for use in similar projects. We start with the introduction of the coarse-grained/discretized image of the epitaxial growth system that we are implementing and then discuss the individual classes into which we separate the problem. As software development is not the main focus of this work, we aim to keep this section short to provide a basic insight into the ideas behind the implementation. We then introduce the "fast diffusion" (FastDiff) algorithm that we have implemented to reduce the computation time spent on simple free diffusion moves and test if it generates the same behaviour as our implementation without FastDiff. Lastly, we will provide an analysis of the computation time scaling with system size and temperature.

## A.1 IMPLEMENTED GEOMETRY AND FRAMEWORK

### A.1.1 LATTICE COARSE-GRAINING

Here, we discuss what kind of lattice geometry we are implementing in our KMC simulation. We are especially interested in simulating the  $C_{60}$  on  $CaF_2(111)$  system where the first layer lattice geometry can be represented by triangular lattice [Fig. 71 left] with a distance between individual lattice sites of  $a_{CaF_2(111)} = 0.386 \text{ nm}$ . The deposited  $C_{60}$  molecules can grow into compact clusters with very little lattice strain, forming a  $C_{60}(111)$  surface for molecules on the second layer to diffuse/nucleate on. The surface lattice geometry from the second layer onward can therefore be represented by a honeycomb lattice with a distance between nearest lattice sites of  $a_{C_{60}(111)} = 0.577 \text{ nm}$  [Fig. 71 middle]. For a completely accurate representation of the  $C_{60}$  on  $CaF_2(111)$  system, we would have to implement both of these lattice structures (triangular fine-grained in the first layer, honeycomb for higher layers). However, we are implementing the same honeycomb lattice for all layers (including the first one), effectively coarse-graining the lattice structure of the substrate [Fig. 71 right]. This brings with it two main advantages:

- Having a consistent lattice structure for all layers makes the implementation way easier. Especially when it comes to the rate refreshing algorithm (where one has to check the states of neighbouring lattice sites), having



**Figure 71:** Coarse-graining of the surface lattice. The fine-grained lattice of the first layer substrate - like the triangular lattice of the  $\text{CaF}_2(111)$  surface - is replaced by a coarse-grained honeycomb lattice which represents the structure of a  $\text{C}_{60}(111)$  surface.

two different lattices in which the particles cover a different amount of lattice sites would add complexity.

- The free diffusion process on the coarse-grained honeycomb lattice needs less individual diffusive transitions to progress, speeding up the simulation.

On the other hand, we are losing the ability to reproduce some of the system's properties:

- The fine-grained diffusion trajectories.
- The order that the specific substrate lattice geometry imposes on the growing clusters. E.g. on  $\text{CaF}_2(111)$ ,  $\text{C}_{60}$  clusters were shown to grow with two distinct orientations [36].
- During coalescence (when a monolayer starts to fill), clusters that grow with different orientations or on mismatching lattice positions are not able to cleanly merge, leaving a grain boundary. This property highly depends on the first layer lattice geometry and is discussed in chapter III for the example of  $\text{C}_{60}$  on mica.

However, the observables that we are interested in for the  $\text{C}_{60}$  on  $\text{CaF}_2(111)$  system are mainly cluster densities (in the early growth regime) and cluster morphologies, which are largely unaffected by the missing properties mentioned above.

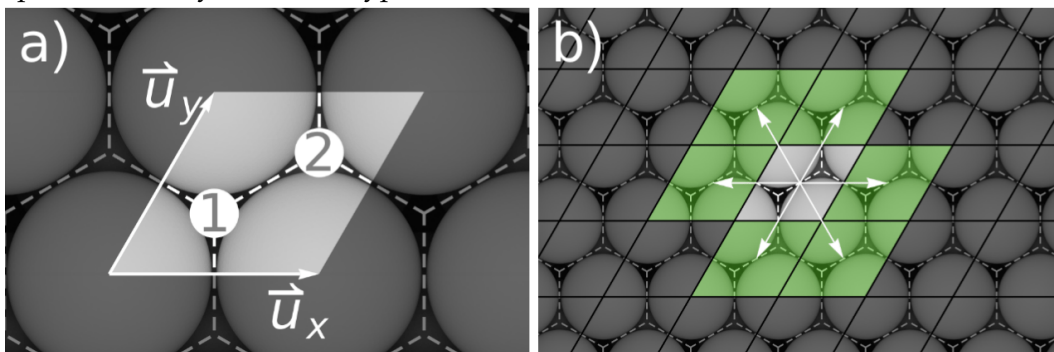
### A.1.2 THE GridNode CLASS

The GridNode class is designed to be a building block for the lattice structure of the simulation and represents a unit cell of the honeycomb lattice [Fig. 72a] that can be replicated in

$$\vec{u}_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} nm, \quad \vec{u}_y = \begin{pmatrix} \cos \frac{\pi}{3} \\ \sin \frac{\pi}{3} \\ 0 \end{pmatrix} nm, \quad \vec{u}_z = \begin{pmatrix} 0 \\ 0 \\ \sqrt{1 - (0.5/\cos \frac{\pi}{6})^2} \end{pmatrix} nm \tag{A.1}$$

directions to build a triangular grid of GridNode objects. This triangular grid is turned into a representation of a honeycomb lattice with the ability for a GridNode to hold particles in distinct sublattice positions  $s$  [white numbered discs in Fig. 72a]. The variable  $s$  can take values  $0$  (GridNode is empty),  $s \in \{1, 2\}$  for an occupied GridNode in the bottom layer ( $z = 0$ ) of the simulation and  $s \in \{s_0 + 1, s_0 + 2\}$  on higher layers ( $z > 0$ ,  $s_0$  is the sublattice position of the supporting cluster of the layer below). In this representation, the positions of particles in the system are encoded in the  $(x, y, z)$  indices of the GridNode

**Figure 72:** Visualization of the GridNode class and some of its components. (a) A GridNode object represents a unit cell that is replicated in  $\vec{u}_x$  and  $\vec{u}_y$  directions to form the honeycomb lattice. An adparticle can take one of the two sublattice positions marked by white dots. (b) GridNode objects are supplied with pointers to neighbouring nodes (green) for easy access to their states. (c) Every GridNode object contains a rate array for double-precision transition rates with dedicated spaces for every transition type and direction.



c) Rate Array (floats with double precision)

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
Diff.		Edge Diff.			Fast Diff.			Ascension			Descension			Overhang As.			Overhang Des.																					

it occupies and its sublattice position  $s$ . From the coordinates  $(x, y, z, s)$ , the "real" coordinates of the particle can be calculated via

$$\begin{pmatrix} x_{\text{real}} \\ y_{\text{real}} \\ z_{\text{real}} \end{pmatrix} = (x + s/3) \cdot \vec{u}_x + (y + s/3) \cdot \vec{u}_y + z \cdot \vec{u}_z. \quad (\text{A.2})$$

In addition to the setup of the lattice geometry, the `GridNode` class also incorporates functions for the determination and storage of transition rates. It is supplied with pointers to neighbouring `GridNode` objects [Fig. 72b] to be able to easily check the environment of an occupying particle and determine its transition rates. To store the determined transition rates, each `GridNode` contains a rate array [Fig. 72c] that can hold transition rates (with double precision) for each of the possible transition types and transition directions.

### A.1.3 THE `RateModel` CLASS

As the transition rates of the particles in the system can only take a finite amount of different values (depending on the type of transition  $i$ , its initial neighbours  $n$ ) at any given temperature  $T$ , it makes sense to calculate these possible transition rate values beforehand and store them in a dedicated space for as long as the temperature of the system stays constant. This is the major purpose of the `RateModel` class. It holds (double precision) values for all the possible transition rates of the system and is also the space where the rate models are implemented (functions that can be used to calculate the transition rate values). The transition rates of the `RateModel` class are accessed by the `GridNode` objects for their transition rate refreshing procedures and the chosen rate model of the system is called to recalculate the transition rates whenever the temperature of the system is changed.

## A.1.4 THE Input CLASS

The Input class is designed to handle the input parameters and instructions supplied by an input file that can look like the example below.

---

```
[PARAMETERS]
BoxDimensions      500 500 2
RandomSeed         Auto
FastDiffusion      Auto
Output             Compressed      trajectory
RateModel          CaF2AtomisticMD 42

[INSTRUCTIONS]
SetTemp            100
SetFlux            0.026 MLPerMin
InsertHexagonalCluster 250 250 0 1 2
Run                MLs 0.1 0.01
SetFlux            0 MLPerMin
SetTemp            240
Run                Time 120 10
```

---

It is separated into a `PARAMETERS` section and an `INSTRUCTIONS` section. The `PARAMETERS` section contains the fixed options that are needed for the KMC simulation to start, like the size of the simulation box, the random seed for the random number generator and the rate model with which the `RateModel` class will calculate the transition rates. The `INSTRUCTIONS` section contains a list of instructions that the program will execute sequentially. Example instructions are changes of temperature or molecular flux (`SetTemp/SetFlux`) or the simulation of a certain amount of time ("`Run Time 60 2`" for simulation of 60 seconds with output to the trajectory file every 2 seconds). The complete set of implemented instructions can be found in the manual on the GitLab repository. With the functionality of the Input class, the program can be used to simulate a wide variety of experimental protocols without the need for recompilation of the source code.

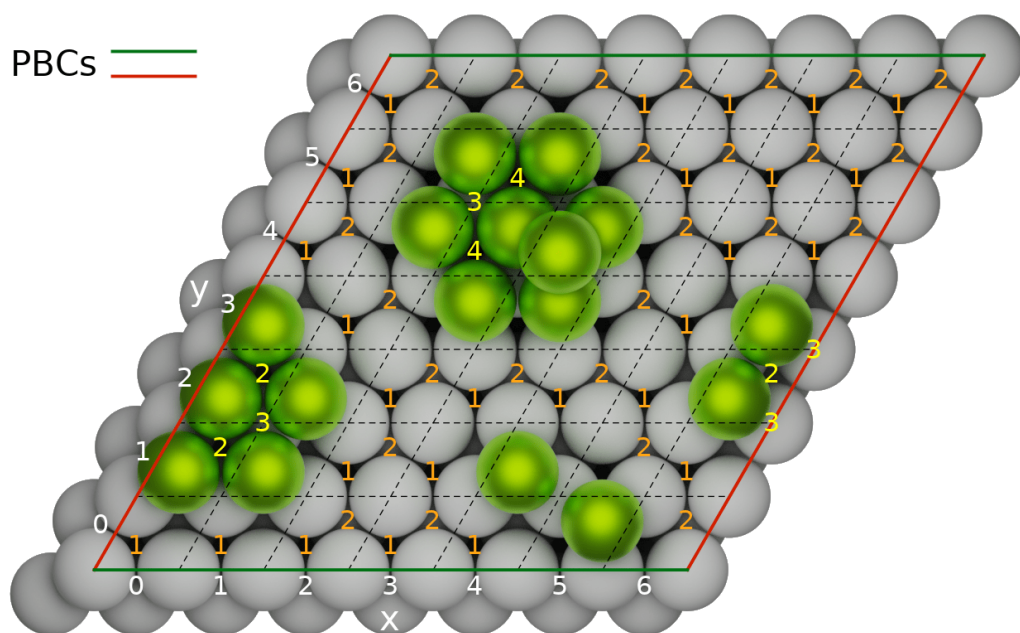


### A.1.5 THE FullereneKMC CLASS

The FullereneKMC class incorporates all the needed functionality to combine the aforementioned classes into a working KMC simulation. To be initialized it needs a path to an input file which is supplied to an Input class to read the system parameters. With the supplied parameters it sets up a random number generator for use in the KMC algorithm, a RateModel to calculate the transition rates and a grid of GridNodes as well as the connections between them [Fig. 72b] including the periodic boundary conditions. It can then execute the instructions given by the Input class.

It is supplied with functions to deposit and move particles around on the grid and to invoke a refreshing of the transition rates of GridNodes in the vicinity to where it makes changes. It keeps track of the time of the system and the total transition rate  $k_{\text{tot}}$  (used to determine the random time steps and the next transition to occur as described in Sec. II.3). To be able to find the chosen next transition of the KMC algorithm without having to go through every single rate array of the GridNode objects [Fig. 72c], the FullereneKMC class is supplied with arrays to store partial sums. These partial sum arrays are based on the system geometry. Let  $k_i(x, y, z), i \in \{0, 1, \dots, 38\}$  be the transition rates saved in

**Figure 73:** Visualization of an example state of the FullereneKMC class. Shown are two clusters and two free adparticles on different sublattice positions in a two-layered system. White numerals are the  $x$  and  $y$  indices of the  $7 \times 7$  GridNodes System. Colored Numerals are the available sublattice positions in the first (orange) and second (yellow) layers. Periodic boundary conditions (PBCs) are implemented in  $x$  and  $y$  direction.



the rate array of `GridNode` at position  $(x,y,z)$ , we start on the bottom level with an array that stores the total transition rate of every `GridNode`,

$$k(x, y, z) = \sum_{i=1}^{38} k_i(x, y, z). \quad (\text{A.3})$$

The next two levels are then given by sums over the all the transition rates of a specific  $(y,z)/z$  coordinate:

$$k(y, z) = \sum_{x=0}^{x_{\max}} k(x, y, z), \quad k(z) = \sum_{y=0}^{y_{\max}} k(y, z). \quad (\text{A.4})$$

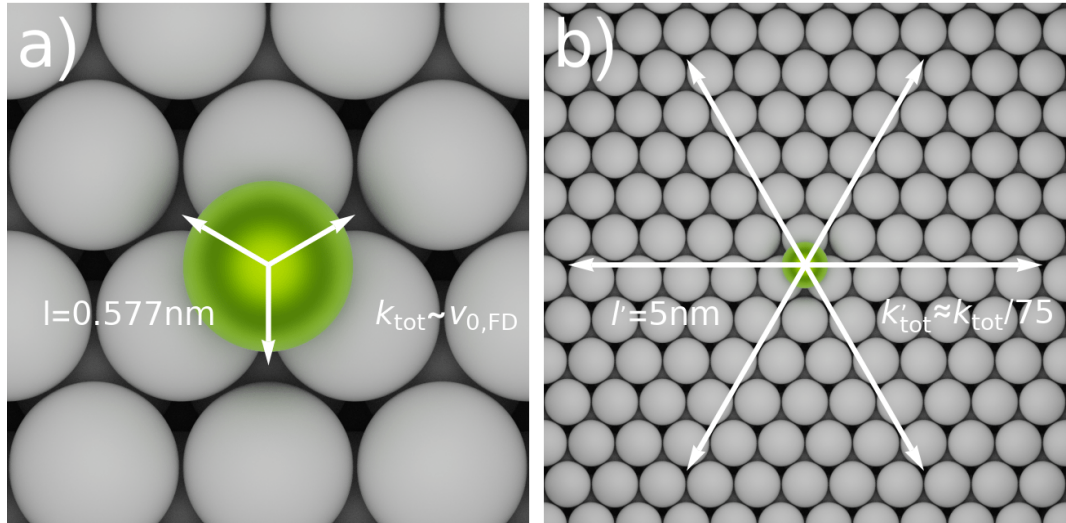
With this, the search algorithm can sequentially find the  $z'$ ,  $y'$  and  $x'$  coordinates of the next transition using the  $k(z)$ ,  $k(y, z')$  and  $k(x, y', z')$  rate sum arrays before checking which exact transition to execute via the rate array of the chosen `GridNode`. The rate sum arrays have to be maintained whenever `GridNodes` recalculate their transition rates. Completely recalculating them would be very expensive, therefore whenever a `GridNode` refreshes their transition rates, it subtracts its overall transition rate value from the rate sum arrays and then adds its new total value after the refreshing. This can and will result in numerical errors due to the addition/subtraction of floating-point variables with various exponents. To mitigate the propagation of those errors, the rate sum arrays are recalculated every few hundred thousand KMC steps and whenever the overall transition rate of the system changes by several orders of magnitude.

## A.2 FAST DIFFUSION

When trying to achieve the experimental system parameters (temperatures of  $T \approx 300$  K and flat terraces of area  $A > 4 \mu\text{m}^2$ ), we observe that most of the steps in the KMC simulation are free diffusion moves, sometimes by fractions of more than 99% (depending on the exact temperature and rate model parameters). As we are not interested in the fine details of the free diffusion trajectories, but rather in the kinetics of the cluster relaxation, our code implements a feature to coarse-grain the free diffusion process for a speed-up of the computation, called "Fast Diffusion" (`FastDiff`).

The idea behind `FastDiff` is based on the formula for the diffusion coefficient,

$$D = \frac{1}{4} l^2 k_{\text{FD}} = \frac{1}{4} l^2 \nu_{0,\text{FD}} e^{-\frac{\Delta E_{\text{FD}}}{k_{\text{B}} T}}, \quad (\text{A.5})$$



**Figure 74:** Example of coarse-graining the diffusion process (a) from a jump length of  $l = 0.577 \text{ nm}$  (b) to a jump length of  $l' = 5 \text{ nm}$ . As a result, the total transition rate  $k_{\text{tot}}$  is reduced by a factor of 75.

with the energy barrier  $\Delta E_{\text{FD}}$ , attempt rate  $\nu_{0,\text{FD}}$  and jump length  $l$  of the free diffusion transition. For any given set of these parameters, based on Eq. A.5 we can transform the jump length and attempt rate via

$$l \rightarrow l' = \alpha \cdot l \quad \nu_{0,\text{FD}} \rightarrow \nu'_{0,\text{FD}} = \frac{\nu_{0,\text{FD}}}{\alpha^2} \quad \alpha \in \mathbb{R}^+ \quad (\text{A.6})$$

while conserving the diffusion coefficient. Using the transformation of Eq. A.6, we can reduce the attempt rates (and thereby the transition rates) of the free diffusion transition significantly by increasing the jump length  $l$  [Fig. 74]. In a rate catalogue that is initially dominated by free diffusion transitions, this results in a significant reduction of the total transition rate  $k_{\text{tot}}$ , thereby speeding up the propagation of the simulation by increasing the time steps  $\Delta t \sim 1/k_{\text{tot}}$  of the KMC algorithm.

We have implemented this idea into our KMC algorithm the following way: Whenever a particle is moved by the KMC algorithm, the immediate vicinity (in a hexagon of circumradius  $R = 2 \text{ nm}$ ) around the particle is checked for the presence of other particles. If particles are present, refresh the transition rates of the particle normally (with fine-grained free diffusion moves). If no other particles are present, the FastDiff algorithm is entered to set transition rates for a coarse-grained diffusion move:

1. Draw a random integer jump length  $l' \in \{1, 2, \dots, l_{\text{max}}\} \text{ nm}$  for the coarse-graining of the next fast diffusion jump of the current particle. The maximum jump length  $l_{\text{max}}$  is a free parameter of the implementation.

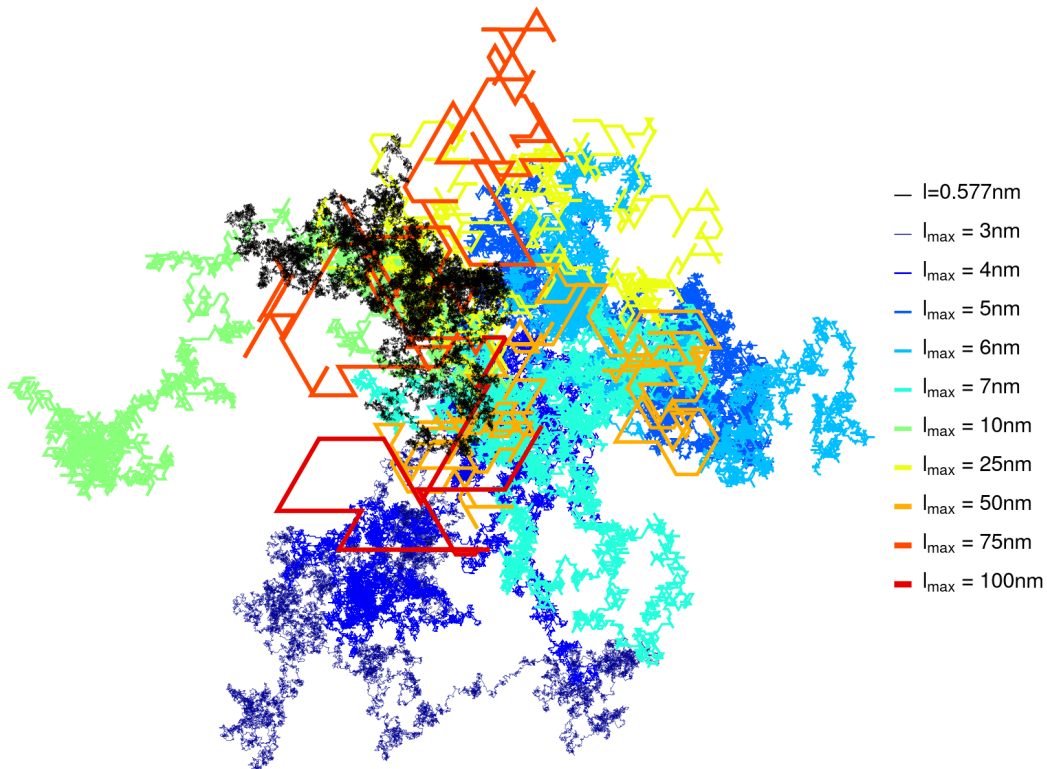
2. Check if any particles are present in a hexagon of circumradius  $l'$ , starting the search from the particle outwards. If a particle is found on the perimeter of a hexagon of circumradius  $l_{\text{part}}$ , set the jump length to  $l' = l_{\text{part}} - 2 \text{ nm}$ .

3. Set the transition rates for the six possible FastDiff directions to

$$\frac{k_{\text{FastDiff},l'}}{6} = \frac{k_{\text{FD}}}{6} \cdot \left( \frac{0.577 \text{ nm}}{l'} \right)^2$$

The added overhead of this algorithm encompasses the need to check a greater vicinity of the particle to set the FastDiff transition rates, more memory to store the additional FastDiff transition rates, and the need to keep track of the fast diffusing particles to be able to detect if any changes of the system make the determined FastDiff transitions invalid (e.g. if a new particle is deposited in the vicinity of a FastDiff particle with a long jump length, requiring a redrawing of the FastDiff jump length). A few example trajectories resulting from a range of different maximum jump lengths  $l_{\text{max}}$  are shown in Fig. 75. In the following we present a few tests of the FastDiff algorithm, comparing the results with normal (fine-grained) free diffusion trajectories.

**Figure 75:** Normal free diffusion trajectory (black) of a single particle in comparison with FastDiff trajectories with equal length in time for several values of  $l_{\text{max}}$ .



### Single particle diffusion test

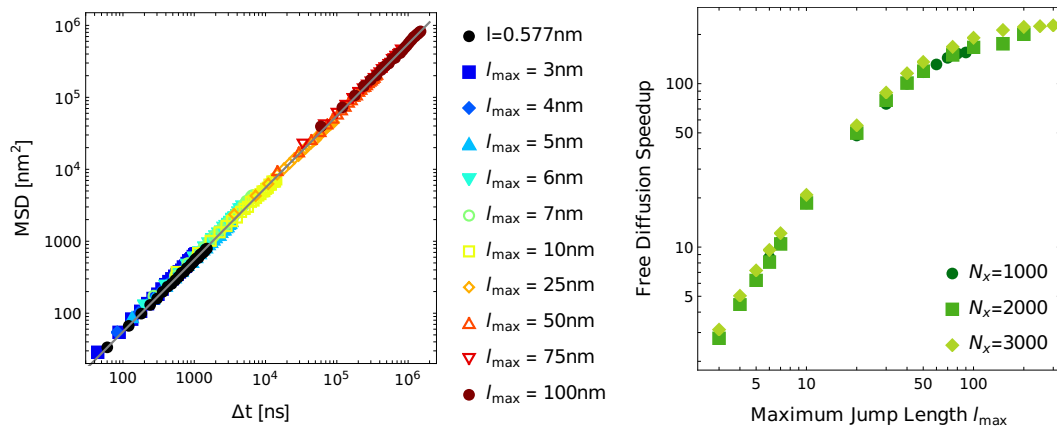
We start with the mean-squared displacement  $\text{MSD}(\Delta t)$  of a freely diffusing particle, which is related to the diffusion coefficient via

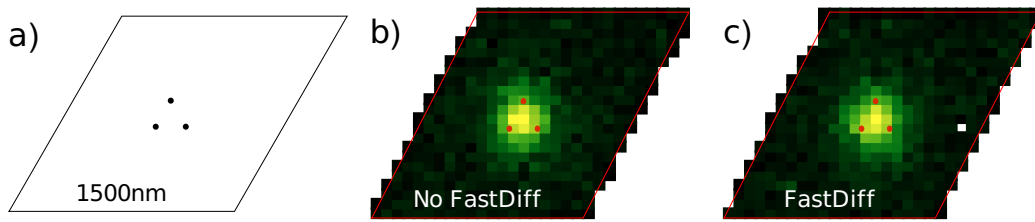
$$\text{MSD}(\Delta t) = \langle x(t + \Delta t) - x(t) \rangle = 4D\Delta t. \quad (\text{A.7})$$

As the diffusion coefficient  $D$  should be conserved under the transformations of Eq. A.6, the slope of the MSD should not be affected by our FastDiff implementation. We measure trajectories of 100000 KMC steps in a  $1000 \times 1000$  unit cell system with a free diffusion barrier of  $\Delta E_{\text{FD}} = 150 \text{ meV}$ , an attempt rate of  $\nu_{0,\text{FD}} = 1 \times 10^{13} \text{ Hz}$  and a temperature of  $T = 200 \text{ K}$ . The obtained MSDs from the trajectories are plotted in Fig. 76 (left), showing that the slope of the MSDs is indeed not affected by the FastDiff implementation. The plot also shows how the 100000 KMC steps of the simulation cover different time and length scales due to the coarse-graining of the free diffusion process.

To test how much faster the KMC simulation progresses with the FastDiff implementation, we set up systems of sizes  $1000 \times 1000$ ,  $2000 \times 2000$  and  $3000 \times 3000$  unit cells with the same diffusion parameters as before and let the simulation run until the trajectory reaches a time of 50 ms and compare the computation time to the simulation without the FastDiff implementation. The obtained speedups are plotted in Fig. 76 (right), showing that a significant speedup of the simulation can be achieved by increasing the maximum jump length  $l_{\text{max}}$ . However, diminishing returns seem to set in for  $l_{\text{max}} \gtrsim 25 \text{ nm}$  and we suggest to not go beyond a maximum value of  $l_{\text{max}} = 75 \text{ nm}$  (with a speedup of roughly a factor of 150).

**Figure 76:** Results of the single-particle diffusion test. Left: The measured MSDs of the generated trajectories. Right: Speedup gained with the FastDiff implementation for single-particle diffusion.



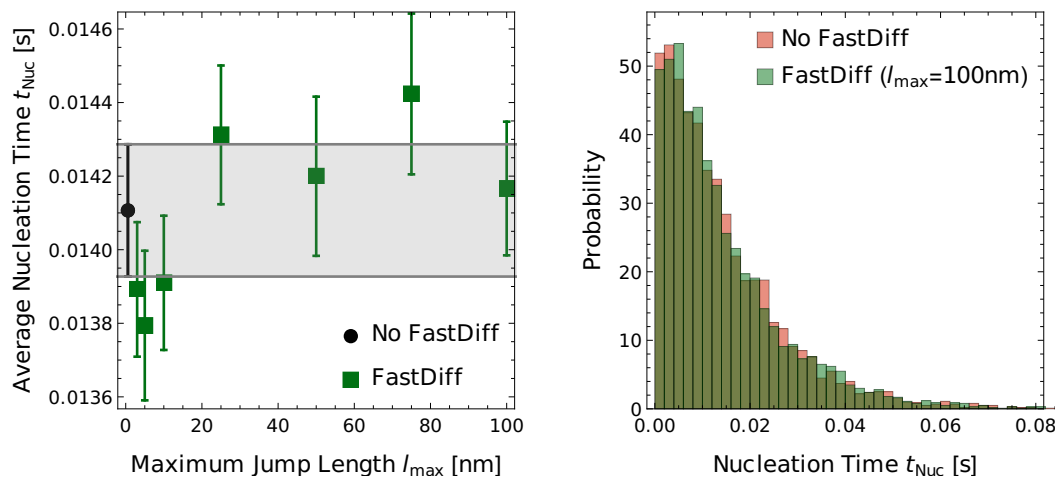


**Figure 77:** Fast Diffusion nucleation test. a) Setup of the three particles with a distance of 200 nm between each other (particles have a diameter of 1 nm and are enlarged to be visible). b) Positional distribution of the nucleated cluster location without the FastDiff and c) with the FastDiff algorithm (at a maximum jump length of  $l_{\max} = 100$  nm).

### Nucleation test

As a test for the collective behaviour of multiple fast diffusing particles, we look at the nucleation of three particles. To this end, we set up a  $1500 \times 1500$  unit cell system and arrange three particles in an even-sided triangle with a distance of 200 nm between each other [Fig. 77a]. As a rate model we use a "hit-and-stick" model in which all transition rates besides free diffusion (again calculated with  $\Delta E_{FD} = 150$  meV,  $\nu_{0,FD} = 1 \times 10^{13}$  Hz and  $T = 200$  K) are set to zero, such that the simulation comes to a halt as soon as the three particles have nucleated. After nucleation, the position of the cluster, as well as the time it took to nucleate (simulation as well as computation time), are output for analysis. For the normal free diffusion and each of the FastDiff parameters ( $l_{\max} \in \{3, 5, 10, 25, 50, 75, 100\}$  nm), 5000 nucleation trajectories are gathered.

**Figure 78:** Nucleation times with normal free diffusion in comparison to the results of the FastDiff implementation. Left: Mean nucleation times. Right: Nucleation time histogram in comparison to the FastDiff algorithm with  $l_{\max} = 100$  nm.



The resulting positional distributions of the nucleated clusters are shown in Fig. 77(b,c) for the normal free diffusion and the FastDiff algorithm (with  $l_{\max} = 100 \text{ nm}$ ), respectively, both showing a similar expected behaviour of most probable nucleation in the middle of the initial triangle. The resulting mean nucleation times are plotted in Fig. 78 (left), showing no significant deviations of the FastDiff implementation results from the fine-grained nucleation trajectories. For a comparison of the nucleation time distributions, histograms of the nucleation times without FastDiff and with FastDiff (at  $l_{\max} = 100 \text{ nm}$ ) are shown in Fig. 78 (right), which also nicely align with each other.

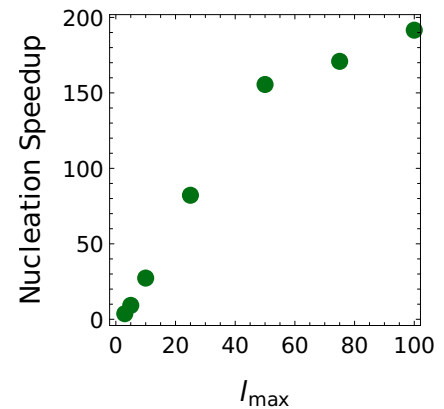


Figure 79: Computation time speedup of the nucleation test.

The achieved speedup of the computation time in this set of nucleation simulations are plotted in Fig. 79. Similar to the previous test, a speedup of a factor 150 – 200 was obtained for maximum jump lengths  $l_{\max} \geq 50 \text{ nm}$  with an observation of diminishing returns above  $l_{\max} = 25 \text{ nm}$ .

#### Options of the FastDiff Implementation

The FastDiff implementation can be controlled with the "FastDiffusion" parameter of the input file. It can be set to "Off" to disable FastDiff, can be supplied with fixed maximum jump lengths via "Fixed lMaxFL lMaxHL" for the first and higher layers respectively, or it can be set to "Auto" to automatically calculate maximum jump lengths based on the other transition rates of the system. The conditions used to determine the maximum jump length in the "Auto" setting are mostly arbitrary:

- We don't want particles to be frozen after deposition, so the total fast diffusion rate per particle should be somewhat larger than the deposition rate. We require  $k_{\text{FastDiff}, l_{\max}} > 100k_{\text{Depos}}$ .
- We don't want the fast diffusion rate to be significantly suppressed in comparison to the common cluster diffusion transitions. For this we require  $k_{\text{FastDiff}, l_{\max}} > 0.1k_{\text{ED-A}}(n = 2)$ .
- We allow  $l_{\max}$  to be maximum 10% of the system dimensions (in x and y directions,  $l_{\max} \leq N_{x/y}$ ). Based on the observed diminishing returns, we set a global maximum value of  $l_{\max} \leq 75 \text{ nm}$ .

As the transition rates of the system can change during the simulation by a change of particle flux or temperature, the maximum jump length will be recalculated in those cases if the "Auto" setting is used.

## A.3 PERFORMANCE

While one is usually interested in simulating a certain experimental protocol with a specific length in time, it is hard to predict the computation time that is going to be needed in the KMC simulation. Here, we are going to show how the computation time can scale with the system's size and also how it strongly varies with the system's temperature in ways that one might not expect. As an example, we use the atomistic rate model determined in Chap. IV with a parameter  $\epsilon_F = 42 \text{ meV}$  and few adjustments:

- The free diffusion rate is calculated with  $\Delta E_{FD} = 187 \text{ meV}$  and  $\nu_{0,FD} = 17 \times 10^{12} \text{ Hz}$ . These values reproduce more closely the experimentally observed cluster densities at low temperatures, and also (more importantly) generate a more pronounced temperature dependence of the computation time.
- The dissociation rate is lowered by a factor of 5 to more closely reproduce the cluster densities at around room temperature.

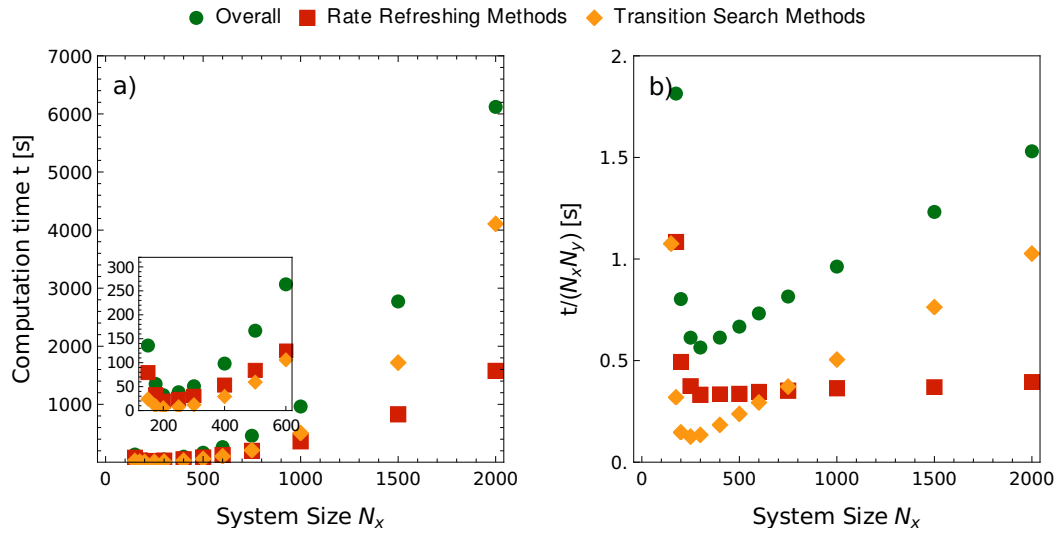
In its current implementation, the RAM usage of the simulation can be a bottleneck as it needs roughly 0.5 kB per unit cell. For the following simulations, we, therefore, do not exceed a system size of  $2000 \times 2000 \times 3$ , taking roughly 6 GB of RAM, which most modern machines can deliver.

### A.3.1 SYSTEM SIZE SCALING

To provide an example for the scaling of the computation time with the system's size, we run simulations of a four minute deposition phase (with a particle flux of  $F = 0.026 \text{ ML/min}$ ) at a constant temperature of  $T = 120 \text{ K}$ . The system is composed with an equal number of unit cells in  $x$  and  $y$  direction ( $N_y = N_x$ ), which is varied in the range  $N_x \in [150, 2000]$ .

The resulting computation times are plotted in Figs. 80(a,b). Subfigure (b) nicely shows how the computation time per unit cell linearly increases with the system size and that this linear increase is mostly generated by the increased search time needed in the transition search method. This scaling can potentially be improved with the implementation of a true binary search tree to enable simulation of even bigger system sizes with little drawback. In its current form, it is more efficient to simulate smaller systems, but one has to be aware that finite-size effects may occur when the system size is so small that only a few clusters are forming in the simulation box. In this example such an effect can be seen for  $N_x < 250$  in Figs. 80(a,b), as the computation time (especially per unit cell) suddenly starts to increase with smaller system sizes.



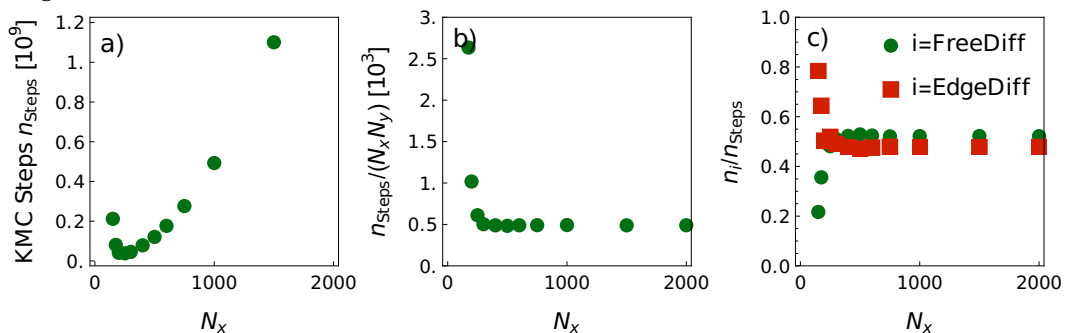


**Figure 80:** Scaling of the overall computation time (green circles) with the system size, as well as the time spent in the two main parts of the algorithm (red squares for rate refreshing, orange diamonds for transition search). (a) total computation time. The inset shows a zoomed-in section of the plot. (b) Computation time per unit cell.

The effect can be explained by the decrease of the total particle deposition rate (as it scales with the substrate area) to a point where nucleated particles have to wait for prolonged periods for the arrival of new particles to form stable configurations. This is reflected in Figs. 81(a,b) with an increase of the total amount of executed KMC steps (as well as the amount of executed KMC steps per unit cell). In Fig. 81c, we can see that the additional KMC steps are mostly of the edge diffusion type, which are executed while the nucleated particles are waiting for the deposition of new particles onto the substrate.

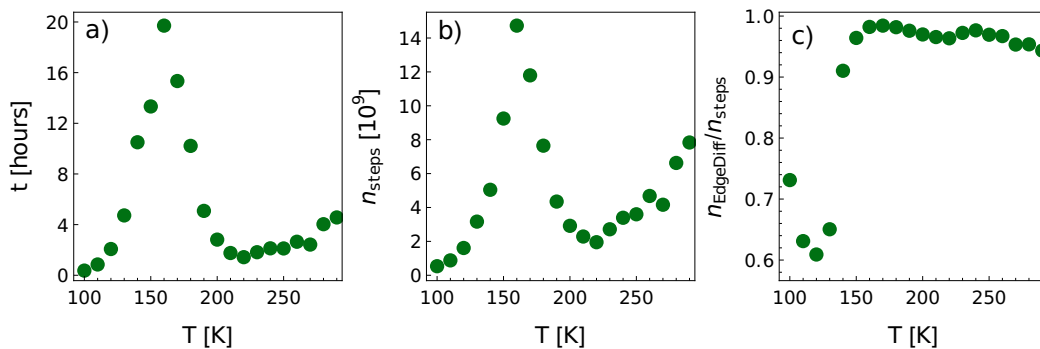
While the system size scaling may change with the system's temperature or with the used rate model, a general takeaway from this section is that one

**Figure 81:** Executed KMC steps versus system size. (a) Total amount of KMC steps executed. (b) KMC steps per unit cell. (c) Proportion of free diffusion and edge diffusion moves executed.



should aim to simulate system sizes in which multiple clusters (optimally in the order of 10 – 100 clusters for best efficiency) can form.

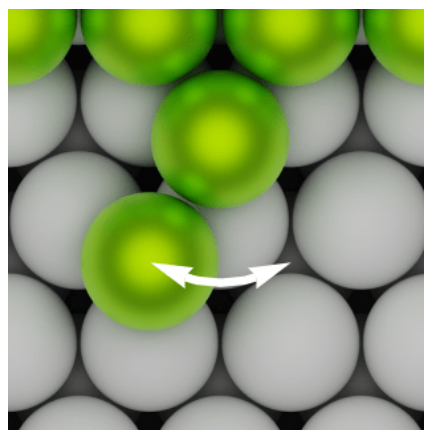
### A.3.2 TEMPERATURE SCALING



**Figure 82:** Temperature scaling of the (a) computation time, (b) the number of executed KMC steps and (c) the proportion of executed edge diffusion moves.

To analyze the scaling of the computation time with temperature, we run simulations with a constant system size of  $1500 \times 1500 \times 3$  unit cells at temperatures in the range  $T \in [100, 290]$  K. The simulation protocol consists of a four minute deposition phase with a flux of  $F = 0.026$  ML/min followed by a four minute post-deposition phase ( $F = 0$ ) in which the clusters can relax.

The results are shown in Figs. 82(a-c). While one might naively expect a general increase of the computation time with the temperature (as the total transition rate increases, the average time step decreases), we observe a large peak of the computation time at a temperature of  $T_{\text{peak}} \approx 160$  K. In this temperature range, the clustered particles are mobile enough to explore many of the edge states and tend to get trapped in one-neighbour edge diffusion states (especially above the first layer where the B-step barrier is very large), in which they do a large number of back-and-forth A-step edge diffusion transitions before finally leaving the state via dissociation or B-step edge diffusion [Fig. 83]. This is reflected in the observed maximum of the proportion of executed edge diffusion transitions in Fig. 82c. With increased temperatures, the decay rate of those states increases faster than the A-step edge diffusion transition rate, leading to a decrease in the computation time spent in those states.



**Figure 83:** "Trapped" state as discussed in the text.

Additionally, the automatically calculated FastDiff jump length increases with temperature, which strongly reduces the overall computation time at higher temperatures.

While the observed behaviour is a specific property of the used rate model, it nicely shows how unpredictable the computation times of a KMC algorithm can be. Especially when changing parameters or features of a rate model, the computation times can change drastically and one should always run a few test simulations before running larger sets of simulations with specified time limits.



# B

## APPENDIX

### CONTENTS

---

B.1	Pseudorandom Number Generators . . . . .	<b>132</b>
B.2	Quaternion approach to obtain optimal rotations . . . . .	<b>133</b>
B.3	Properties of the Exponential Distribution . . . . .	<b>135</b>
B.4	Example LAMMPS input script . . . . .	<b>140</b>
B.5	Acknowledgements . . . . .	<b>161</b>

---

## B.1 PSEUDORANDOM NUMBER GENERATORS

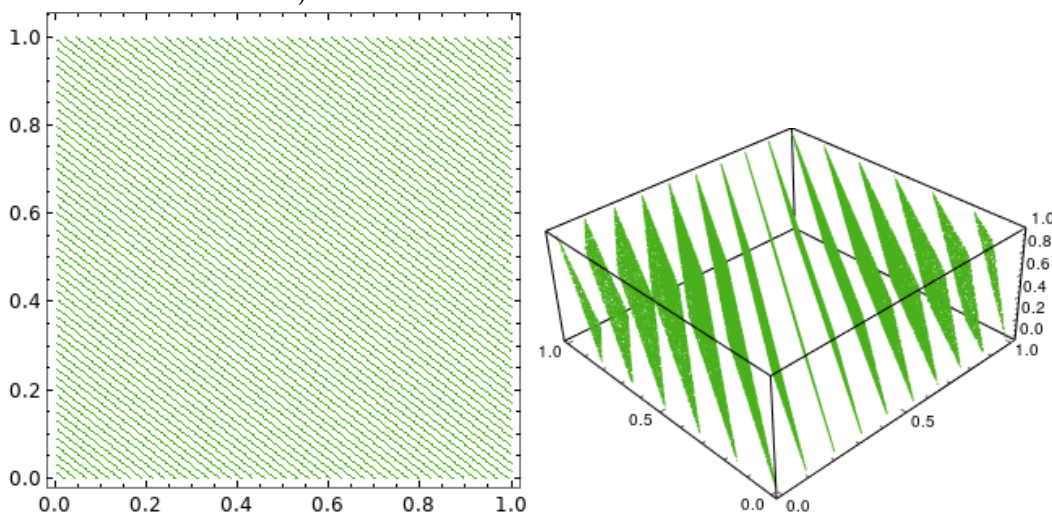
Monte Carlo methods are dependent on a source of random numbers. While there are ways to obtain "true" random numbers, e.g. through measurement of some physical observable that is understood as being random (e.g. atmospheric noise [164]), one usually resorts to the use of so-called pseudorandom number generators (PRNGs) as they are easier and more efficient to use. PRNGs can take an initial "seed" value and generate a string of values based on some mathematical algorithm. These strings of numbers are in principle deterministic (as a PRNG always generates the same string if supplied with the same seed), however, a PRNG can aim to produce strings whose properties approximate the properties of real strings of random numbers.

Probably the simplest class of PRNGs are linear congruential generators (LCGs), which were widely used for a major part of the 20th century before better algorithms were developed. LCGs are based on the recurrence relation

$$x_{n+1} = (a \cdot x_n + b) \pmod{c} \quad (\text{B.1})$$

with a modulus  $c$  ( $c > 0$ ), a multiplier  $a$  ( $0 < a < c$ ) and an increment  $b$  ( $0 < b < m$ ). The LCG is supplied with a seed value  $x_0$  from which it can then derive a string of random numbers  $\{x_1, x_2, \dots\}$ . While LCGs are still widely used because of their simplicity and efficiency (mostly for simple applications), they come with a few weaknesses that make them undesirable for scientific

**Figure 84:** Hyperplanes found in the output of an LCGs. Left: In two dimensions from the output of an LCG with  $a = 69069, b = 0, c = 2^{16}$ . Right: In three dimensions from the output of an LCG with  $a = 65539, b = 0, c = 2^{31}$  (the so-called "RANDU" LCG).



applications like MC simulations. An example of those weaknesses is shown in Fig. 84: when used to choose coordinates in an  $n$ -dimensional space, the generated points tend to fall onto a fixed number of hyperplanes, separated by areas that are never realised.

More sophisticated PRNGs have been developed to avoid many of the ills of LCGs, like the Mersenne Twister [165], which is the algorithm that we have chosen to use in our KMC simulation. Specifically we make use of the "MT19937" implementation of the GNU Scientific Library [166].

## B.2 QUATERNION APPROACH TO OBTAIN OPTIMAL ROTATIONS

For the determination of the rotational diffusion in Sec. IV.1.2, we make use of a quaternion approach [158, 159] to determine optimal rotation transformations between two conformations of the  $C_{60}$  molecule at different times of the trajectory. Here, we outline the necessary calculations for our specific example. A conformation  $A$  of the  $C_{60}$  molecule is determined by the  $3 \cdot 60$  coordinates of its carbon atoms (the origin of the coordinate frame is the center of mass of the  $C_{60}$  molecule) and can be written as

$$A = \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ \dots & \dots & \dots \\ x_{60} & y_{60} & z_{60} \end{pmatrix}, \quad \left[ A^T = \begin{pmatrix} x_1 & x_2 & \dots & x_{60} \\ y_1 & y_2 & \dots & y_{60} \\ z_1 & z_2 & \dots & z_{60} \end{pmatrix} \right], \quad (\text{B.2})$$

where  $A^T$  is the transpose of the conformation matrix. To work towards a optimal rotation between two conformations at different times,  $A(t_1)$  and  $A(t_2)$ , we start by calculating the inner product  $M$ ,

$$M = A^T(t_1)A(t_2) = \begin{pmatrix} S_{xx} & S_{xy} & S_{xz} \\ S_{yx} & S_{yy} & S_{yz} \\ S_{zx} & S_{zy} & S_{zz} \end{pmatrix}, \quad (\text{B.3})$$

where the matrix components are defined like  $S_{xy} := \sum_{i=1}^{60} x_i(t_1)y_i(t_2)$ . The optimal rotation matrix can then be determined in its unit quaternion represen-

tation as the eigenvector  $\mathbf{q}$  with the most positive eigenvalue of the symmetric  $4 \times 4$  matrix  $\mathbf{K}^{10}$  derived from  $\mathbf{M}$  via

$$\mathbf{K}^{10} = \begin{pmatrix} S_{xx} + S_{yy} + S_{zz} & S_{yz} - S_{zy} & S_{zx} - S_{xz} & S_{xy} - S_{yx} \\ S_{yz} - S_{zy} & S_{xx} - S_{yy} - S_{zz} & S_{xy} + S_{yx} & S_{zx} + S_{xz} \\ S_{zx} - S_{xz} & S_{xy} + S_{yx} & S_{yy} - S_{xx} - S_{zz} & S_{yz} + S_{zy} \\ S_{xy} - S_{yx} & S_{zx} + S_{xz} & S_{yz} + S_{zy} & S_{zz} - S_{xx} - S_{yy} \end{pmatrix}. \quad (\text{B.4})$$

The quaternion  $\mathbf{q}$  can be written as

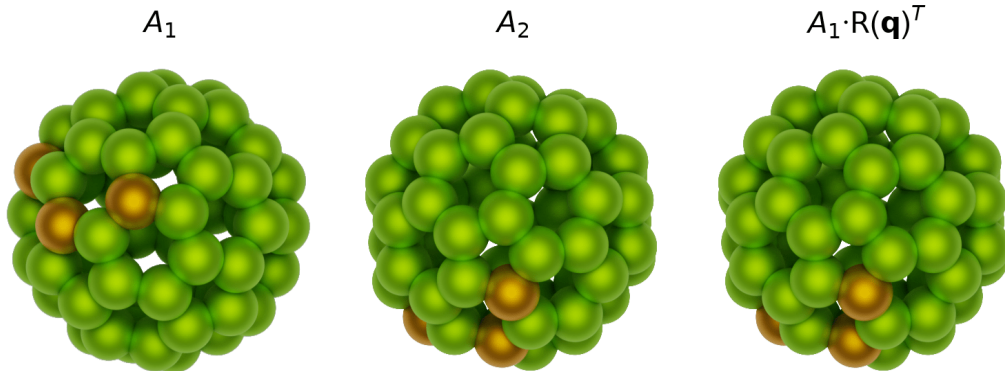
$$\mathbf{q} = \begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} \cos(\theta/2) \\ \vec{v} \sin(\theta/2) \end{pmatrix} \quad |\vec{v}| = 1 \quad (\text{B.5})$$

and represents a rotation of an angle  $\theta$  about a (3-dimensional) rotation axis  $\vec{v}$ . For our analysis of mean-squared angular displacements (MSAD), we take the values  $\theta = 2 \cdot \arccos(q_0)$  as the angular displacement between the two conformations. The quaternion  $\mathbf{q}$  can be translated into a rotation matrix  $\mathbf{R}(\mathbf{q})$  via

$$\mathbf{R}(\mathbf{q}) = \begin{pmatrix} 1 - 2q_2^2 - 2q_3^2 & 2q_1q_2 - 2q_0q_3 & 2q_1q_3 + 2q_0q_2 \\ 2q_2q_1 + 2q_0q_3 & 1 - 2q_3^2 - 2q_1^2 & 2q_2q_3 - 2q_0q_1 \\ 2q_3q_1 - 2q_0q_2 & 2q_3q_2 + 2q_0q_1 & 1 - 2q_1^2 - 2q_2^2 \end{pmatrix}, \quad (\text{B.6})$$

which can be used to do a sanity check if the obtained rotation is valid [Fig. 85].

**Figure 85:** Example conformations  $A_1$  and  $A_2$  as well as the result obtained by rotating conformation  $A_1$  with the optimal rotation  $\mathbf{R}(\mathbf{q})$ . Three atoms are highlighted in orange to be able to track the orientation of the conformations.





## B.3 PROPERTIES OF THE EXPONENTIAL DISTRIBUTION

The waiting times for the processes we are describing in KMC simulations are exponentially distributed. To be able to understand why the KMC algorithm is a valid way of simulating the dynamics of our given system, it is therefore important to understand some of the characteristics of the exponential distribution, which we are going to derive in this section. The derivations of this section can also be found in basic statistics/stochastic textbooks like Ref. [167] and in my master's thesis [86].

### *Probability Distribution of the Waiting Time for the Next Occurring Transition*

Suppose we start with a given system in a microstate which it can leave through  $n$  possible elementary transition pathways. We assume that the set of random exponentially distributed waiting times  $\mathbf{X} := \{X_1, X_2, \dots, X_n\}$  for each of the transitions is a set of independent random variables. We can describe the waiting time  $X_i$  for transition  $i \in \{1, 2, \dots, n\}$  with the exponential distribution function with parameter  $k_i$  via

$$P(X_i \leq t) = 1 - e^{-k_i t} \quad \text{for } t \geq 0.$$

For the KMC algorithm, we are interested in the waiting time until the first of the possible transitions occurs. In terms of random variables, we are interested in the distribution function of the random variable  $Z := \min(\mathbf{X})$ , which we can derive the following way:

$$\begin{aligned} P(Z \leq t) &= P(\min(\mathbf{X}) \leq t) = 1 - P(\min(\mathbf{X}) > t) \\ &= 1 - P(X_1 > t, \dots, X_n > t) \stackrel{(X_i \text{ indep.})}{=} 1 - \prod_{i=1}^n P(X_i > t) \\ &= 1 - \prod_{i=1}^n (1 - P(X_i \leq t)) \\ &= 1 - \prod_{i=1}^n e^{-k_i t} \\ &= 1 - e^{-\sum_i k_i t} = 1 - e^{-k_{\text{tot}} t}. \end{aligned}$$

We find that the random variable  $Z$  is also described by an exponential distribution, the parameter of which is the total rate  $k_{\text{tot}} = \sum_i k_i$ . This finding justifies the way the KMC algorithm advances the time in the simulation by

drawing exponentially distributed time steps according to the total transition rate of the system.

### *Probability of Event $i$ Being the Next Occurring Transition*

Knowing how the waiting time for the next occurring transition is distributed, we are now interested in the probability that any given transition  $i$  is the next one that is going to be executed. In terms of the defined random variables, we want to derive the probability for  $X_i$  to be the smallest waiting time of the set  $\mathbf{X}$ ,

$$P(X_i < X_j \forall j \neq i).$$

To derive this probability we are making use of the probability densities of our random variables:

$$p_{X_i}(t) = \frac{d}{dt}P(X_i \leq t) = k_i e^{-k_i t}.$$

Probability densities can be used to calculate the probability of a random variable  $X_i$  taking a value in a given interval  $(x, y)$  via

$$P(x < X_i < y) = \int_x^y p_{X_i}(t) dt.$$

As the random variables  $X_i$  are independent, we can write down the probability density of the whole set  $\mathbf{X}$  as the product of the individual probability densities

$$p_{X_1, \dots, X_n}(t_1, \dots, t_n) = \prod_{i=1}^n k_i e^{-k_i t_i}.$$

We can now use this distribution to calculate the probability for every variable  $X_j, j \neq i$  to be larger than  $X_i$ ,

$$\begin{aligned} P(X_i < X_j \forall j \neq i) &= \int_0^\infty k_i e^{-k_i t_i} \left( \prod_{j \neq i} \int_{t_i}^\infty k_j e^{-k_j t_j} dt_j \right) dt_i \\ &= \int_0^\infty k_i e^{-k_i t_i} \left( \prod_{j \neq i} e^{-k_j t_i} \right) dt_i \\ &= \int_0^\infty k_i e^{-\sum_j k_j t_i} dt_i = \int_0^\infty k_i e^{-k_{\text{tot}} t_i} dt_i \\ &= \frac{k_i}{k_{\text{tot}}}. \end{aligned}$$

We find that the probability of event  $i$  to have the smallest waiting time (i.e. for it to be the next transition to occur) is explicitly given by the ratio of its transition rate  $k_i$  with the total transition rate  $k_{\text{tot}}$ . This probability is reproduced by the KMC algorithms way of determining the next transition via the drawing of a random number  $r \in [0, k_{\text{tot}}]$  and the sequential search for the first partial rate sum  $s_m = \sum_{j=1}^m k_j$ ,  $m \in \{1, 2, \dots, n\}$  that satisfies  $s_m < r$  (As described in Sec. II.3).

### *Memorylessness of the Exponential Distribution*

Whenever we do a KMC simulation step by choosing a time step  $\Delta t$  and a transition  $i$  to execute according to the probabilities derived in the previous sections, we end up in a new state in which some of the previously possible transitions may become unrealisable (e.g. because of blocked diffusion paths) while other new transitions can become available. I.e. for the possible transitions of the next KMC step, we are left with a subset  $\mathbf{X}' \subseteq \mathbf{X}$  of "old" transitions – which have already waited a time  $\Delta t$  – and we have a set  $\mathbf{X}^*$  of "new" random variables for the transitions that were just made available. The question of how we can account for the fact that some of our transitions have already waited a while for their possible occurrence is luckily very easy to answer for exponentially distributed waiting times. We simply have to look at the conditional probability distribution of the "old" random variables  $\mathbf{x}'$  given that we know that they have already waited a time  $\Delta t$ :

$$\begin{aligned}
 P(X'_j \leq t + \Delta t | X'_j > \Delta t) &= 1 - P(X'_j > t + \Delta t | X'_j > \Delta t) \\
 &= 1 - \frac{P(X'_j > t + \Delta t \cap X'_j > \Delta t)}{P(X'_j > \Delta t)} \\
 &\stackrel{t \geq 0}{=} 1 - \frac{P(X'_j > t + \Delta t)}{P(X'_j > \Delta t)} \\
 &= 1 - \frac{1 - P(X'_j \leq t + \Delta t)}{1 - P(X'_j \leq \Delta t)} \\
 &= 1 - \frac{e^{-k_j(t+\Delta t)}}{e^{-k_j\Delta t}} = 1 - e^{-k_j t}.
 \end{aligned}$$

We find that the conditional probability distribution of the "old" transitions is exactly the same as the ones of a "new" exponentially distributed variable. This property is the memorylessness of the exponential distribution. It simplifies the implementation of KMC algorithms as it allows equal treatment of all possible

transitions, independent of the fact that some may have already waited some time for their occurrence during previous KMC steps.

*Determining the transition rates for the multiple possible outgoing transitions of a given initial state*

Suppose we are looking at an initial state that has multiple possible escape pathways that we would ascribe to different types of transition  $i$  [Fig. 86]. We assume that each transition  $i$  has a waiting time  $X_i$  that is an exponentially distributed random variable (with rate  $k_i$ ) that is independent of the other possible transitions. The probability distribution for such a waiting time  $X_i$  can be written as

$$P(X_i \leq t) = 1 - e^{-k_i t} \quad \text{for } t \geq 0. \quad (\text{B.7})$$

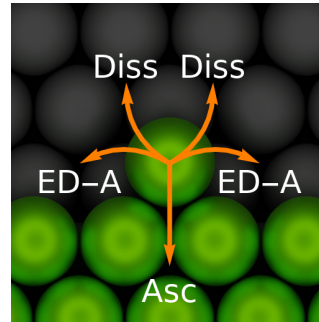


Figure 86: Example configuration.

So for an initial state with  $n$  possible transitions, we have a set  $\mathbf{X} = \{X_1, \dots, X_n\}$  of independent exponentially distributed random variables to describe the process. The particle will escape the initial state after the time has reached the lowest of the waiting times ( $t_e = \min(\mathbf{X}) = X_m$ ), and it will move to the state that is the target of transition  $m$ . In our simulations, we are measuring the escape time from the initial state, and take a snapshot of the state of the system after the transition occurred (which tells us the type of transition). A set of  $N$  simulations will yield  $N$  escape times  $t_{e,j}$  as well as a number of total occurrences for each transition,  $N_i$  ( $N = \sum_i N_i$ ). We need to derive informations about these observables to be able to analyze the data correctly.

First, we take a look at the escape time  $t_e$ , which is always going to be the minimum of the waiting times  $\{X_1, \dots, X_n\}$ . The probability distribution for  $t_e$  can be derived as (see earlier section)

$$\begin{aligned} P(t_e \leq t) &= P(\min(\{X_1, \dots, X_n\}) \leq t) = 1 - e^{-\sum_i k_i t} \\ &= 1 - e^{k_{\text{tot}} t}, \end{aligned}$$

so it behaves like a exponentially distributed random variable with total rate  $k_{\text{tot}} = \sum_i k_i$ . From our set of  $N$  simulations with escape times  $t_{e,j}$  we can therefore calculate  $k_{\text{tot}}$  with the maximum likelihood estimator

$$\boxed{\hat{k}_{\text{tot}} = \frac{N}{t_{\text{tot}}}} \quad \text{or unbiased} \quad \boxed{\hat{k}_{\text{tot}} = \frac{N-1}{t_{\text{tot}}}} \quad \left( t_{\text{tot}} = \sum_j^N t_{e,j} \right). \quad (\text{B.8})$$

Next, for each individual simulation, the probability of transition  $i$  to occur can be written as: (see earlier section)

$$P(X_i < X_j \forall j \neq i) = \frac{k_i}{k_{\text{tot}}}.$$

So if we are interested in the occurrence of transition  $i$ , we can view each simulation as a Bernoulli experiment with success chance  $p = k_i/k_{\text{tot}}$ . The total number of occurrences  $n_i$  over the course of  $N$  simulations is therefore described by the binomial distribution

$$P(n_i \times \text{transition } i) = b_{N, \frac{k_i}{k_{\text{tot}}}}(n_i) \quad (\text{B.9})$$

$$= \binom{N}{n_i} \left(\frac{k_i}{k_{\text{tot}}}\right)^{n_i} \left(1 - \frac{k_i}{k_{\text{tot}}}\right)^{N-n_i}. \quad (\text{B.10})$$

The maximum likelihood estimator for  $k_i$  can be derived by requiring

$$\frac{d}{dk_i} b_{N, \frac{k_i}{k_{\text{tot}}}}(n_i) = 0 \quad \rightarrow \quad \boxed{\hat{k}_i = \frac{n_i}{N} \cdot k_{\text{tot}}}, \quad (\text{B.11})$$

where  $k_{\text{tot}}$  can either be replaced by one of the estimators in Eq. (B.8) or by an estimation derived from the distribution of escape times (as shown in Sec. II.3.3).

## B.4 EXAMPLE LAMMPS INPUT SCRIPT

Below is listed one of the input scripts we have used for one of the earliest simulations of a single  $C_{60}$  molecule freely diffusing on a  $C_{60}(111)$  surface (here at  $T = 300$  K). As an interaction potential, a tabulated Girifalco potential was supplied via the file `girifalco.table`. Comprehensive explanations for all the commands can be found in the online LAMMPS documentation at <https://lammps.sandia.gov/doc/>.

```
#example surface diffusion script for C60 on C60

#basic system parameters
dimension      3
units          metal
atom_style     atomic
boundary       p p f

#definition of a grid on which to place particles
lattice custom 1.0 &
               a1 10.0 0.0 0.0 &
               a2 5.0 8.66 0.0 &
               a3 5.0 2.86 8.14 &
               basis 0.2 0.2 0.2

#creation of the simulation box
region         box block 0 4 0 6 0 10
create_box     2 box

#creation of particles
region         substrate block INF INF INF INF INF 4
create_atoms   1 region substrate
create_atoms   2 single 3.00 2.0 4.5

#Definition of pair potentials
pair_style     table linear 1200
pair_coeff     * * girifalco.table GIRIFALCO_C60 20
pair_coeff     1 2 girifalco.table GIRIFALCO_C60 20
mass          * 720.66
neigh_modify   delay 0

#grouping of particles
group         addatoms type 2
```

```
region          mobile block 0 4 0 6 1 4
group           mobile region mobile
region          slab block 0 3 0 5 5 6

# compute commands to access temperature of mobile particles
compute         mob mobile temp
compute_modify  mob dynamic yes extra 0

#Some log output options
thermo_style    custom step atoms temp epair etotal press
thermo          10000
thermo_modify   temp mob

#Randomization Run
fix             1 addatoms nve #nve: Verlet integrator
fix             3 mobile nve
fix             2 mobile langevin 300 300 0.01 332115279
timestep        0.01
run             100000

#Temp Monitoring Run
variable        newTemp equal "300/f_2*c_mob"
reset_timestep  0
unfix           2
fix             2 mobile ave/time 50 2 200 c_mob ave window 250
fix             4 mobile temp/rescale 50000 v_newTemp 300 1.0 1.0
run             352000

#Measurement Run
reset_timestep  0
unfix           4
unfix           2
dump            1 addatoms xyz 200 ./OutputFile.out
run             40000000
```

---





# LIST OF FIGURES

Figure 1	$C_{60}$ deposition experiments on various insulating substrates . . . . .	2
Figure 2	Sketch of a molecular beam epitaxy setup. . . . .	6
Figure 3	Deposition rate of $C_{60}$ from a Knudsen cell. . . . .	7
Figure 4	Sketch of an NC-AFM setup. . . . .	8
Figure 5	$CaF_2(111)$ structure. . . . .	9
Figure 6	Atomistic visualization of $C_{60}$ on $CaF_2(111)$ . . . . .	10
Figure 7	AFM Images from experiments of $C_{60}$ on $CaF_2(111)$ . . . . .	10
Figure 8	Basic elementary diffusion transitions. . . . .	12
Figure 9	Edge Diffusion along A and B Edges. . . . .	14
Figure 10	Ascension to overhang sites. . . . .	14
Figure 11	Ehrlich-Schwöbel barrier. . . . .	15
Figure 12	Cluster shape change: edge diffusion/dissociation. . . . .	16
Figure 13	Cluster shape change: B step edge diffusion. . . . .	16
Figure 14	Cluster shape change: Ascension. . . . .	17
Figure 15	Ascension onto A and B edge positions. . . . .	17
Figure 16	Surface diffusion as an example for an infrequent-event system. . . . .	22
Figure 17	Illustration of the rfKMC pathway selection . . . . .	24
Figure 18	Additive Interactions . . . . .	26
Figure 19	Density plot of the potential energy of a $C_{60}$ molecule in a cluster configuration . . . . .	26
Figure 20	Illustration of TST . . . . .	28
Figure 21	Illustration of the drag method . . . . .	29
Figure 22	Illustration of the NEB method . . . . .	30
Figure 23	Transition rates from MD simulations . . . . .	32
Figure 24	Example probability distribution of residence times. . . . .	32
Figure 25	The Lennard-Jones interaction potential. . . . .	35
Figure 26	Illustration of the velocity Verlet algorithm . . . . .	37
Figure 27	Periodic Boundary Conditions . . . . .	38
Figure 28	Equilibration using velocity rescaling . . . . .	39
Figure 29	$C_{60}$ on $C_{60}(111)$ intro figure . . . . .	44
Figure 30	Girifalco potential derivation sketch . . . . .	45

Figure 31	Girifalco potential comparison with experiment and atomistic potential samples. . . . .	47
Figure 32	$C_{60}$ on $C_{60}(111)$ example configuration (Descension) .	48
Figure 33	Potential energy and particle coordination as stop triggers . . . . .	49
Figure 34	Energy barriers versus the number of initial neighbours	53
Figure 35	Entropy production measurement system. . . . .	56
Figure 36	Entropy production in the RawMD and Simple model	57
Figure 37	Cluster Morphologies in the RawMD and Simple model . . . . .	58
Figure 38	Shape descriptors . . . . .	59
Figure 39	Cluster feature plots for the RawMD and Simple model	60
Figure 40	Cluster densities, experiment versus Simple model . .	62
Figure 41	$C_{60}$ multilayer growth snapshots. . . . .	63
Figure 42	Grain boundaries and the evolution of separated domains on $C_{60}(111)$ and mica . . . . .	64
Figure 43	Peak cluster densities, experiment vs our upper and lower bounds . . . . .	66
Figure 44	Configuration for the C-C interaction potential tuning.	72
Figure 45	Parameter tuning of the atomistic carbon-carbon interaction . . . . .	72
Figure 46	Illustration of the interaction coarse-graining . . . . .	74
Figure 47	Free Diffusion of $C_{60}$ on $CaF_2(111)$ MD configuration	75
Figure 48	MEPs for Surface diffusion . . . . .	77
Figure 49	$CaF_2(111)$ surface unit cell and positional probability distribution . . . . .	78
Figure 50	Free Energies from positional probability distributions	79
Figure 51	Probability density and potential energy plots . . . .	79
Figure 52	Example diffusion trajectories for the atomistic and coarse-grained model . . . . .	80
Figure 53	Arrhenius plots of translational and rotational diffusion coefficients . . . . .	81
Figure 54	Example MSAD plots . . . . .	82
Figure 55	Summary of obtained Free Diffusion Barriers on $CaF_2(111)$ . . . . .	84
Figure 56	Unconstrained Arrhenius fits versus constrained modelling of the free diffusion coefficients . . . . .	88
Figure 57	Two cluster orientations of $C_{60}$ on $CaF_2(111)$ . . . . .	89
Figure 58	Initial configuration for edge diffusion with one initial neighbour. Particles of interest are highlighted in orange. . . . .	90

Figure 59	Unconstrained Arrhenius fits versus constrained modelling of edge diffusion rates ( $n = 1$ ) . . . . .	91
Figure 60	Initial configurations for edge diffusion with two initial neighbours. Particles that can go through transitions of interest are highlighted in orange. . . . .	92
Figure 61	Unconstrained Arrhenius fits versus constrained modelling of edge diffusion rates ( $n = 2$ ) . . . . .	93
Figure 62	Unconstrained Arrhenius fits versus constrained modelling of ascension rates ( $n = 2$ ) . . . . .	94
Figure 63	Unconstrained Arrhenius fits versus constrained modelling of dissociation rates ( $n = 2$ ) . . . . .	95
Figure 64	Transitions to and from overhang sites . . . . .	98
Figure 65	Cluster density measurements in KMC simulations for $C_{60}$ on $CaF_2(111)$ . . . . .	100
Figure 66	Impact of the critical cluster size $i^*$ on the island densities at higher temperatures . . . . .	102
Figure 67	$\epsilon_F$ parameter tuning via second layer occupation . . . . .	103
Figure 68	Low-temperature morphology comparison . . . . .	105
Figure 69	Room temperature morphology comparison . . . . .	106
Figure 70	Growth trajectories at 310 K starting with a second layer nucleus at different times of the growth process. . . . .	111
Figure 71	KMC lattice coarse-graining . . . . .	115
Figure 72	GridNode Class Visualization . . . . .	116
Figure 73	Visualization of a FullerenesKMC example system . . . . .	119
Figure 74	FastDiff Example . . . . .	121
Figure 75	FastDiff Trajectories . . . . .	122
Figure 76	FastDiff MSD Test . . . . .	123
Figure 77	FastDiff nucleation test setup and position densities. . . . .	124
Figure 78	FastDiff nucleation times. . . . .	124
Figure 79	Computation time speedup of the nucleation test. . . . .	125
Figure 80	Computation time vs. system size . . . . .	127
Figure 81	Number of KMC steps vs. system size . . . . .	127
Figure 82	Computation time vs. temperature . . . . .	128
Figure 83	Trapped State . . . . .	128
Figure 84	Hyperplanes of LCGs . . . . .	132
Figure 85	Quaternion approach example result. . . . .	134
Figure 86	Rate determination example configuration. . . . .	138



# LIST OF TABLES

Table 1	$C_{60}$ on $C_{60}(111)$ MD simulation results . . . . .	50
Table 2	Continuation of Tab. 1 . . . . .	51
Table 3	Comparison of some of our $C_{60}$ on $C_{60}(111)$ energy barriers with previous results from the literature . . .	52
Table 4	RawMD model energy barriers . . . . .	53
Table 5	RawMD model attempt rates . . . . .	54
Table 6	Parameters taken from Ref. [150] for the repulsive and Van der Waals interactions of $\phi_{ij}^S$ [Eq. (IV.1)]. . . .	71
Table 7	Compilation of some of the interaction parameters derived in Ref. [151] and the parameters that we have chosen to use for our atomistic model (bottom two lines) for eqs. IV.5 and IV.6. (O: oxygen of carbonate group, Ow: Oxygen of water, CD: carbon of methanoic acid) . . . . .	73
Table 8	MEP analysis results. All values are given in meV . .	76
Table 9	Rate model parameters for the free diffusion process derived from the translational diffusion coefficients. Listed are the parameters of Eq. (IV.24) together with the attempt frequencies (calculated with the assumption of single jumps of length $l = 0.386$ nm). . . . .	87
Table 10	Final model parameters derived from our MD simulations. Listed are the parameters of Eq. (IV.24) together with the attempt frequencies. . . . .	97
Table 11	Optimal $\epsilon_F$ values for the two models and resulting energy barriers, attempt rates and transition rates for ascension with two initial neighbours ( $n = 2$ ). For comparison, the values used by Martin Körner [53] in KMC simulations investigating the same system are also listed. . . . .	103



# BIBLIOGRAPHY

- [1] William Janke and Thomas Speck. *Modeling of epitaxial film growth of C<sub>60</sub> revisited*. In: *Phys. Rev. B* 101.12 (Mar. 2020), p. 125427.
- [2] William Janke et al. *Erratum: Quantitative description of C<sub>60</sub> diffusion on an insulating surface [Phys. Rev. B 82, 155428 (2010)]*. In: *Phys. Rev. B* 101.4 (Jan. 2020), p. 049907.
- [3] William Janke and Thomas Speck. *Multiscale modeling of structure formation of C<sub>60</sub> on insulating CaF<sub>2</sub> substrates*. In: *J. Chem. Phys.* 154.23 (June 2021), p. 234701.
- [4] Michael Graetzel et al. *Materials interface engineering for solution-processed photovoltaics*. In: *Nature* 488.7411 (Aug. 2012), pp. 304–312.
- [5] D. Vuillaume. *Molecular electronics based on self-assembled monolayers*. Ed. by A.V. Narlikar and Y.Y. Fu. Oxford University Press, July 2017.
- [6] R. Hoffmann-Vogel. *Imaging prototypical aromatic molecules on insulating surfaces: a review*. In: *Rep. Prog. Phys.* 81.1 (Dec. 2017), p. 016501.
- [7] Angelika Kühnle. *Self-assembly of organic molecules at metal surfaces*. In: *Curr. Opin. Colloid Interface Sci.* 14.2 (2009), pp. 157–168.
- [8] Johannes V. Barth. *Molecular Architectonic on Metal Surfaces*. In: *Annu. Rev. Phys. Chem.* 58.1 (2007), pp. 375–407.
- [9] Mario Einax, Wolfgang Dieterich, and Philipp Maass. *Colloquium: Cluster growth on surfaces: Densities, size distributions, and morphologies*. In: *Rev. Mod. Phys.* 85 (3 2013), pp. 921–939.
- [10] Philipp Rahe et al. *Tuning Molecular Self-Assembly on Bulk Insulator Surfaces by Anchoring of the Organic Building Blocks*. In: *Adv. Mater.* 25.29 (June 2013), pp. 3948–3956.
- [11] Felix Kling et al. *Self-assembly of Organic Molecules on Insulating Surfaces*. In: *Noncontact Atomic Force Microscopy*. Springer International Publishing, 2015, pp. 147–171.
- [12] Roberto Otero et al. *Molecular Self-Assembly at Solid Surfaces*. In: *Adv. Mater.* 23.44 (Sept. 2011), pp. 5148–5176.

- [13] H. W. Kroto et al. *C<sub>60</sub>: Buckminsterfullerene*. In: *Nature* 318.6042 (Nov. 1985), pp. 162–163.
- [14] M. A. Verheijen et al. *Growth and morphology of C<sub>60</sub> crystals*. In: *Chem. Phys. Lett.* 191.3-4 (Apr. 1992), pp. 339–344.
- [15] J. Li et al. *Growth and morphology of C<sub>60</sub> and C<sub>70</sub> single crystals*. In: *J. Cryst. Growth* 143.1-2 (Oct. 1994), pp. 58–65.
- [16] W. I. F. David et al. *Structural Phase Transitions in the Fullerene C<sub>60</sub>*. In: *Europhys. Lett.* 18.3 (Feb. 1992), pp. 219–225.
- [17] Roger Moret. *Orientalional ordering in the low-temperature phase of C<sub>60</sub> studied by single-crystal x-ray diffraction*. In: *Phys. Rev. B* 48.23 (Dec. 1993), pp. 17619–17621.
- [18] H. Kasatani et al. *X-ray-diffraction study of the phase transition in a C<sub>60</sub> single crystal*. In: *Phys. Rev. B* 47.7 (Feb. 1993), pp. 4022–4024.
- [19] Y. Yoneda, K. Sakaue, and T. Terauchi. *Phase transitions of thin films grown by molecular beam epitaxy*. In: *J. Phys. Condens. Matter* 9.14 (Apr. 1997), pp. 2851–2857.
- [20] Sergey I. Bozhko et al. *Rotational transitions in a C<sub>60</sub> monolayer on the WO<sub>2</sub>/W(110) surface*. In: *Phys. Rev. B* 84.19 (Nov. 2011), p. 195412.
- [21] S. I. Bozhko et al. *Rotation dynamics of C<sub>60</sub> molecules in a monolayer fullerene film on the WO<sub>2</sub>/W(110) surface near the rotational phase transition*. In: *J. Exp. Theor. Phys.* 120.5 (May 2015), pp. 831–837.
- [22] L. A. Girifalco. *Interaction Potential for C<sub>60</sub> Molecules*. In: *J. Phys. Chem.* 95.14 (1991), pp. 5370–5371.
- [23] L. A. Girifalco. *Molecular properties of fullerene in the gas and solid phases*. In: *J. Phys. Chem.* 96.2 (1992), pp. 858–861.
- [24] J. Weckesser, J. V. Barth, and K. Kern. *Mobility and bonding transition of C<sub>60</sub> on Pd(110)*. In: *Phys. Rev. B* 64 (16 2001), p. 161403.
- [25] Song Guo et al. *Thermal Diffusion of C<sub>60</sub> Molecules and Clusters on Au(111)*. In: *J. Phys. Chem. B* 108.37 (Sept. 2004), pp. 14074–14081.
- [26] Felix Loske et al. *Growth of ordered C<sub>60</sub> islands on TiO<sub>2</sub>(110)*. In: *Nanotechnology* 20.6 (Jan. 2009), p. 065606.
- [27] Rémy Pawlak et al. *Atomic-Scale Mechanical Properties of Orientated C<sub>60</sub> Molecules Revealed by Noncontact Atomic Force Microscopy*. In: *ACS Nano* 5.8 (July 2011), pp. 6349–6354.
- [28] R. Pawlak et al. *High-resolution imaging of C<sub>60</sub> molecules using tuning-fork-based non-contact atomic force microscopy*. In: *J. Phys. Condens. Matter* 24.8 (Feb. 2012), p. 084005.



- [29] Andrea Picone et al. *Controlling the Electronic and Structural Coupling of C<sub>60</sub> Nano Films on Fe(001) through Oxygen Adsorption at the Interface*. In: *ACS Appl. Mater. Interfaces* 8.39 (Sept. 2016), pp. 26418–26424.
- [30] Dong Chen and Dror Sarid. *Growth of C<sub>60</sub> films on silicon surfaces*. In: *Surf. Sci.* 318.1-2 (Oct. 1994), pp. 74–82.
- [31] Wei Chen et al. *C<sub>60</sub> on SiC Nanomesh*. In: *J. Phys. Chem. B* 110.43 (Nov. 2006), pp. 21873–21881.
- [32] S. Szuba et al. *Observation of C<sub>60</sub> film formation on a highly oriented pyrolytic graphite substrate via scanning tunnelling microscopy*. In: *Appl. Surf. Sci.* 144-145 (Apr. 1999), pp. 648–652.
- [33] Hui Liu and Petra Reinke. *C<sub>60</sub> thin film growth on graphite: Coexistence of spherical and fractal-dendritic islands*. In: *J. Chem. Phys.* 124.16 (2006), p. 164707.
- [34] Heekeun Shin et al. *Floating two-dimensional solid monolayer C<sub>60</sub> on graphite*. In: *Phys. Rev. B* 82.23 (Dec. 2010), p. 235427.
- [35] S. A. Burke et al. *C<sub>60</sub> on alkali halides: Epitaxy and morphology studied by noncontact AFM*. In: *Phys. Rev. B* 76.3 (July 2007), p. 035419.
- [36] Felix Loske et al. *Quantitative description of C<sub>60</sub> diffusion on an insulating surface*. In: *Phys. Rev. B* 82 (15 2010), p. 155428.
- [37] M. A. Groce et al. *Temperature-dependent nucleation and capture-zone scaling of C<sub>60</sub> on silicon oxide*. In: *Surf. Sci.* 606.1-2 (Jan. 2012), pp. 53–56.
- [38] Philipp Rahe et al. *From dewetting to wetting molecular layers: C<sub>60</sub> on CaCO<sub>3</sub>(1014) as a case study*. In: *Phys. Chem. Chem. Phys.* 14.18 (2012), p. 6544.
- [39] N. V. Sibirev et al. *Size distributions of fullerene surface clusters*. In: *Appl. Surf. Sci.* 307 (July 2014), pp. 46–51.
- [40] Kazuma Sato et al. *Morphological phase diagrams of C<sub>60</sub> and C<sub>70</sub> films on graphite*. In: *Surf. Sci.* 664 (Oct. 2017), pp. 222–225.
- [41] Hiroki Mitsuta et al. *Epitaxial Growth of C<sub>60</sub> on Rubrene Single Crystals for a Highly Ordered Organic Donor/Acceptor Interface*. In: *Cryst. Growth Des.* 17.9 (Aug. 2017), pp. 4622–4627.
- [42] E. Seydel, R. Hoffmann-Vogel, and M. Marz. *Epitaxial growth of C<sub>60</sub> on highly oriented pyrolytic graphite surfaces studied at low temperatures*. In: *Nanotechnology* 30.2 (Nov. 2018), p. 025703.
- [43] Andrea Huttner, Tobias Breuer, and Gregor Witte. *Controlling Interface Morphology and Layer Crystallinity in Organic Heterostructures: Microscopic View on C<sub>60</sub> Island Formation on Pentacene Buffer Layers*. In: *ACS Appl. Mater. Interfaces* 11.38 (Aug. 2019), pp. 35177–35184.

- [44] Nguyen Ngan Nguyen et al. *Charge-Transfer-Controlled Growth of Organic Semiconductor Crystals on Graphene*. In: *Adv. Sci.* 7.6 (Mar. 2020), p. 1902315.
- [45] Lu'an Guo et al. *Orientational Epitaxy of van der Waals Molecular Heterostructures*. In: *Nano Lett.* 18.8 (July 2018), pp. 5257–5261.
- [46] Katsumi Tanigaki et al. *Crystal growth of C<sub>60</sub> thin films on layered substrates*. In: *Appl. Phys. Lett.* 63.17 (Oct. 1993), pp. 2351–2353.
- [47] Frédéric Rossel et al. *Growth and characterization of fullerene nanocrystals on NaCl/Au(111)*. In: *Phys. Rev. B* 84.7 (Aug. 2011), p. 075426.
- [48] A. V. Matetskiy et al. *Peculiar diffusion of C<sub>60</sub> on In-adsorbed Si(111) $\sqrt{3} \times \sqrt{3}$ -Au surface*. In: *Surf. Sci.* 616 (Oct. 2013), pp. 44–50.
- [49] Arthur F. Voter. *Introduction to the Kinetic Monte Carlo Method*. In: *Radiation Effects in Solids*. Springer Netherlands, pp. 1–23.
- [50] Daniel T. Gillespie. *A general method for numerically simulating the stochastic time evolution of coupled chemical reactions*. In: *J. Comput. Phys.* 22.4 (1976), pp. 403–434.
- [51] Daniel T. Gillespie. *Exact stochastic simulation of coupled chemical reactions*. In: *J. Phys. Chem.* 81.25 (1977), pp. 2340–2361.
- [52] Hui Liu et al. *Fractal Structures in Fullerene Layers: Simulation of the Growth Process*. In: *J. Phys. Chem. C* 112.12 (2008), pp. 4687–4695.
- [53] Martin Körner et al. *Second-Layer Induced Island Morphologies in Thin-Film Growth of Fullerenes*. In: *Phys. Rev. Lett.* 107 (1 2011), p. 016101.
- [54] Rebecca A. Cantrell and Paulette Clancy. *A New Kinetic Monte Carlo Algorithm for Heteroepitactical Growth: Case Study of C<sub>60</sub> Growth on Pentacene*. In: *J. Chem. Theory Comput.* 8.3 (Feb. 2012), pp. 1048–1057.
- [55] S. Bommel et al. *Unravelling the multilayer growth of the fullerene C<sub>60</sub> in real time*. In: *Nat. Commun.* 5 (2014), p. 5388.
- [56] Nicola Kleppmann and Sabine H. L. Klapp. *Particle-resolved dynamics during multilayer growth of C<sub>60</sub>*. In: *Phys. Rev. B* 91 (4 2015), p. 045436.
- [57] N. Kleppmann, F. Schreiber, and S. H. L. Klapp. *Limits of size scalability of diffusion and growth: Atoms versus molecules versus colloids*. In: *Phys. Rev. E* 95 (2 2017), p. 020801.
- [58] Yaset M. Acevedo et al. *Multiscale Simulation and Modeling of Multilayer Heteroepitactic Growth of C<sub>60</sub> on Pentacene*. In: *Langmuir* 32.12 (2016), pp. 3045–3056.
- [59] Simon Aeschlimann et al. *Focus on the Essential: Extracting the Decisive Energy Barrier of a Complex Process*. In: *Adv. Mater. Interfaces* 6.20 (2019), p. 1900795.

- [60] Felix Loske. "Tailoring molecule nanostructures on insulating surfaces investigated by non-contact atomic force microscopy". PhD thesis. 2011.
- [61] S. A. Burke, J. M. Topple, and P. Grütter. *Molecular dewetting on insulators*. In: *J. Phys. Condens. Matter* 21.42 (Sept. 2009), p. 423101.
- [62] W. Patrick McCray. *MBE deserves a place in the history books*. In: *Nature Nanotechnology* 2.5 (May 2007), pp. 259–261.
- [63] John E. Davey and Titus Pankey. *Epitaxial GaAs Films Deposited by Vacuum Evaporation*. In: *J. Appl. Phys.* 39.4 (Mar. 1968), pp. 1941–1948.
- [64] A. Y. Cho and J. R. Arthur. *Molecular beam epitaxy*. In: *Prog. Solid. State Ch.* 10 (Jan. 1975), pp. 157–191.
- [65] Louis Royer. *Recherches expérimentales sur l'épitaxie ou orientation mutuelle de cristaux d'espèces différentes*. In: *Bull. Soc. Fr. Min.* 51.1 (1928), pp. 7–159.
- [66] Thomas Michely and Joachim Krug. *Introduction*. In: *Islands, Mounds and Atoms*. Springer Berlin Heidelberg, 2004, pp. 1–11.
- [67] G. Binnig et al. *Tunneling through a controllable vacuum gap*. In: *Appl. Phys. Lett.* 40.2 (Jan. 1982), pp. 178–180.
- [68] G. Binnig et al. *7×7 Reconstruction on Si(111) Resolved in Real Space*. In: *Phys. Rev. Lett.* 50.2 (Jan. 1983), pp. 120–123.
- [69] G. Binnig, C. F. Quate, and Ch. Gerber. *Atomic Force Microscope*. In: *Phys. Rev. Lett.* 56.9 (Mar. 1986), pp. 930–933.
- [70] T. R. Albrecht et al. *Frequency modulation detection using high-Q cantilevers for enhanced force microscope sensitivity*. In: *J. Appl. Phys.* 69.2 (Jan. 1991), pp. 668–673.
- [71] F. J. Giessibl. *Atomic Resolution of the Silicon (111)-(7×7) Surface by Atomic Force Microscopy*. In: *Science* 267.5194 (Jan. 1995), pp. 68–71.
- [72] Hitoshi Ueyama et al. *Atomically Resolved InP(110) Surface Observed with Noncontact Ultrahigh Vacuum Atomic Force Microscope*. In: *Jap. J. Appl. Phys.* 34.Part 2, No. 8B (Aug. 1995), pp. L1086–L1088.
- [73] F. J. Giessibl. *Subatomic Features on the Silicon (111)-(7x7) Surface Observed by Atomic Force Microscopy*. In: *Science* 289.5478 (July 2000), pp. 422–425.
- [74] M. Emmrich et al. *Subatomic resolution force microscopy reveals internal structure and adsorption sites of small iron clusters*. In: *Science* 348.6232 (Mar. 2015), pp. 308–311.
- [75] A. M. Sweetman et al. *Mapping the force field of a hydrogen-bonded assembly*. In: *Nat. Commun.* 5.1 (May 2014).

- [76] C. Chiutu et al. *Precise Orientation of a Single C<sub>60</sub> Molecule on the Tip of a Scanning Probe Microscope*. In: *Phys. Rev. Lett.* 108 (26 2012), p. 268302.
- [77] S. Morita, R. Wiesendanger, and E. Meyer, eds. *Noncontact Atomic Force Microscopy*. Springer Berlin Heidelberg, 2002.
- [78] S. Morita, F. J. Giessibl, and R. Wiesendanger, eds. *Noncontact Atomic Force Microscopy*. Vol. 2. Springer Berlin Heidelberg, 2009.
- [79] S. Morita et al., eds. *Noncontact Atomic Force Microscopy*. Vol. 3. Springer International Publishing, 2015.
- [80] M. Reichling et al. *Degradation of the CaF<sub>2</sub>(111) surface by air exposure*. In: *Surf. Sci.* 439.1-3 (Sept. 1999), pp. 181–190.
- [81] L. Tröger et al. *Concept for support and cleavage of brittle crystals*. In: *Rev. Sci. Instrum.* 80.6 (June 2009), p. 063703.
- [82] J. A. Venables. *Rate equation approaches to thin film nucleation kinetics*. In: *Philos. Mag.* 27.3 (Mar. 1973), pp. 697–738.
- [83] J. A. Venables, G. D. T. Spiller, and M. Hanbucken. *Nucleation and growth of thin films*. In: *Rep. Prog. Phys.* 47.4 (Apr. 1984), pp. 399–459.
- [84] Svante Arrhenius. *Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren*. In: *J. Phys. Chem.* 4 (Jan. 1889), p. 226.
- [85] Thomas Michely and Joachim Krug. *Condensation, Diffusion and Nucleation*. In: *Islands, Mounds and Atoms*. Springer Berlin Heidelberg, 2004, pp. 13–59.
- [86] William Janke. “Cluster Growth of Organic Molecules”. Unpublished master’s thesis. Johannes Gutenberg-Universität Mainz, 2017.
- [87] D. Walton. *Nucleation of Vapor Deposits*. In: *J. Chem. Phys.* 37.10 (Nov. 1962), pp. 2182–2188.
- [88] C.-M Zhang et al. *Submonolayer island formation and the onset of multilayer growth during Ag/Ag(100) homoepitaxy*. In: *Surf. Sci.* 406.1-3 (May 1998), pp. 178–193.
- [89] Nicholas Metropolis et al. *Equation of State Calculations by Fast Computing Machines*. In: *J. Chem. Phys.* 21.6 (1953), pp. 1087–1092.
- [90] Abhijit Chatterjee and Dionisios G. Vlachos. *An overview of spatial microscopic and accelerated kinetic Monte Carlo methods*. In: *J. Comput. Aided Mat. Des.* 14.2 (2007), pp. 253–308.
- [91] A. B. Bortz, M. H. Kalos, and J. L. Lebowitz. *A new algorithm for Monte Carlo simulation of Ising spin systems*. In: *J. Comput. Phys.* 17.1 (1975), pp. 10–18.

- [92] Mark F. Gyure et al. *Unstable Growth on Rough Surfaces*. In: *Phys. Rev. Lett.* 81.22 (Nov. 1998), pp. 4931–4934.
- [93] Aleksy K. Jones et al. *Faceting at the step flow threshold in epitaxial growth on patterned surfaces*. In: *Phys. Rev. B* 79.20 (May 2009).
- [94] I. K. Marmorkos and S. Das Sarma. *Atomistic numerical study of molecular-beam-epitaxial growth kinetics*. In: *Phys. Rev. B* 45.19 (May 1992), pp. 11262–11272.
- [95] M. R. Marcelin. *Contribution à l'étude de la cinétique physico-chimique*. In: *Ann. Physique* 9.3 (1915), pp. 120–231.
- [96] Henry Eyring. *The Activated Complex in Chemical Reactions*. In: *J. Chem. Phys.* 3 (1935), p. 107.
- [97] M. G. Evans and M. Polanyi. *Some applications of the transition state method to the calculation of reaction velocities, especially in solution*. In: *Trans. Faraday Soc.* 31 (1935), p. 875.
- [98] George H. Vineyard. *Frequency factors and isotope effects in solid state rate processes*. In: *J. Phys. Chem. Solids* 3.1 (1957), pp. 121–127.
- [99] G. DeLorenzi, C. P. Flynn, and G. Jacucci. *Effect of anharmonicity on diffusion jump rates*. In: *Phys. Rev. B* 30 (10 1984), pp. 5430–5448.
- [100] Mads R. So/rensen and Arthur F. Voter. *Temperature-accelerated dynamics for simulation of infrequent events*. In: *J. Chem. Phys.* 112.21 (2000), pp. 9599–9606.
- [101] Graeme Henkelman, Gísli Jóhannesson, and Hannes Jónsson. *Methods for Finding Saddle Points and Minimum Energy Paths*. In: *Theoretical Methods in Condensed Phase Chemistry*. Ed. by Steven D. Schwartz. Dordrecht: Springer Netherlands, 2002, pp. 269–302.
- [102] Hannes Jónsson, Greg Mills, and Karsten W. Jacobsen. *Nudged elastic band method for finding minimum energy paths of transitions*. In: *Classical and Quantum Dynamics in Condensed Phase Simulations*. World Scientific, June 1998.
- [103] Greg Mills and Hannes Jónsson. *Quantum and thermal effects in H<sub>2</sub> dissociative adsorption: Evaluation of free energy barriers in multidimensional quantum systems*. In: *Phys. Rev. Lett.* 72.7 (Feb. 1994), pp. 1124–1127.
- [104] Gregory Mills, Hannes Jónsson, and Gregory K. Schenter. *Reversible work transition state theory: application to dissociative adsorption of hydrogen*. In: *Surf. Sci.* 324.2-3 (Feb. 1995), pp. 305–337.
- [105] Thomas A. Halgren and William N. Lipscomb. *The synchronous-transit method for determining reaction pathways and locating molecular transition states*. In: *Chem. Phys. Lett.* 49.2 (July 1977), pp. 225–232.

- [106] Michael J. Rothman and Lawrence L. Lohr. *Analysis of an energy minimization method for locating transition states on potential energy hypersurfaces*. In: *Chem. Phys. Lett.* 70.2 (Mar. 1980), pp. 405–409.
- [107] Marie Villarba and Hannes Jónsson. *Diffusion mechanisms relevant to metal crystal growth: Pt/Pt(111)*. In: *Surf. Sci.* 317.1-2 (Sept. 1994), pp. 15–36.
- [108] Marie Villarba and Hannes Jónsson. *Atomic exchange processes in sputter deposition of Pt on Pt(111)*. In: *Surf. Sci.* 324.1 (Feb. 1995), pp. 35–46.
- [109] Mads R. Sørensen, Karsten W. Jacobsen, and Hannes Jónsson. *Thermal Diffusion Processes in Metal-Tip-Surface Interactions: Contact Formation and Adatom Mobility*. In: *Phys. Rev. Lett.* 77.25 (Dec. 1996), pp. 5067–5070.
- [110] T. Rasmussen et al. *Atomistic Determination of Cross-Slip Pathway and Energetics*. In: *Phys. Rev. Lett.* 79.19 (Nov. 1997), pp. 3676–3679.
- [111] Blas P. Uberuaga et al. *Diffusion of Ge below the Si(100) Surface: Theory and Experiment*. In: *Phys. Rev. Lett.* 84.11 (Mar. 2000), pp. 2441–2444.
- [112] Graeme Henkelman, Blas P. Uberuaga, and Hannes Jónsson. *A climbing image nudged elastic band method for finding saddle points and minimum energy paths*. In: *J. Chem. Phys.* 113.22 (Dec. 2000), pp. 9901–9904.
- [113] Isaac Newton. *Philosophiæ naturalis principia mathematica*. Jussu Societatis Regiæ ac Typis Josephi Streater. Prostat apud plures bibliopolas, 1687.
- [114] E. Fermi et al. *Studies of Nonlinear Problems*. Tech. rep. May 1955.
- [115] B. J. Alder and T. E. Wainwright. *Studies in Molecular Dynamics. I. General Method*. In: *J. Chem. Phys.* 31.2 (Aug. 1959), pp. 459–466.
- [116] J. B. Gibson et al. *Dynamics of Radiation Damage*. In: *Phys. Rev.* 120.4 (Nov. 1960), pp. 1229–1253.
- [117] A. Rahman. *Correlations in the Motion of Atoms in Liquid Argon*. In: *Phys. Rev.* 136.2A (Oct. 1964), A405–A411.
- [118] Steve Plimpton. *Fast Parallel Algorithms for Short-Range Molecular Dynamics*. In: *J. Comput. Phys.* 117.1 (1995), pp. 1–19. (LAMMPS Project website: <http://lammps.sandia.gov>).
- [119] D. Andrienko P. Virnau G. Settanni. *Lecture notes in computersimulations in statistical physics*. 2014/2015.
- [120] Michael P. Allen and Dominic J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, Nov. 2017.

- [121] Daan Frenkel and Berend Smit. *Molecular Dynamics Simulations*. In: *Understanding Molecular Simulation*. Elsevier, 2002, pp. 63–107.
- [122] J. E. Lennard-Jones. *Cohesion*. In: *Proc. Phys. Soc.* 43.5 (Sept. 1931), pp. 461–482.
- [123] C. A. Coulomb. *Mémoires sur l'électricité et la magnétisme*. Chez Bachelier, libraire, 1789.
- [124] William C. Swope et al. *A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters*. In: *J. Chem. Phys.* 76.1 (Jan. 1982), pp. 637–649.
- [125] Daan Frenkel and Berend Smit. *Long-Range Interactions*. In: *Understanding Molecular Simulation*. Elsevier, 2002, pp. 291–320.
- [126] M. H. J. Hagen et al. *Does C<sub>60</sub> have a liquid phase?* In: *Nature* 365.6445 (Sept. 1993), pp. 425–426.
- [127] C. Caccamo, D. Costa, and A. Fucile. *A Gibbs ensemble Monte Carlo study of phase coexistence in model C<sub>60</sub>*. In: *J. Chem. Phys.* 106.1 (Jan. 1997), pp. 255–263.
- [128] M. Hasegawa and K. Ohno. *Monte Carlo simulation study of the high-temperature phase diagram of model C<sub>60</sub> molecules*. In: *J. Chem. Phys.* 111.13 (Oct. 1999), pp. 5955–5963.
- [129] Rui P. S. Fartaria, Fernando M. S. Silva Fernandes, and Filomena F. M. Freitas. *Monte Carlo Simulation of the Phase Diagram of C<sub>60</sub> Using Two Interaction Potentials. Enthalpies of Sublimation*. In: *J. Phys. Chem. B* 106.39 (Oct. 2002), pp. 10227–10232.
- [130] Ailan Cheng, Michael L. Klein, and Carlo Caccamo. *Prediction of the phase diagram of rigid C<sub>60</sub> molecules*. In: *Phys. Rev. Lett.* 71.8 (Aug. 1993), pp. 1200–1203.
- [131] K. Kniaż et al. *Fullerene alloys*. In: *Solid State Comm.* 96.10 (Dec. 1995), pp. 739–743.
- [132] R. M. Khusnutdinoff, A. V. Mokshin, and I. D. Takhaviev. *Microscopic dynamics of an amorphous C<sub>60</sub><sub>x</sub>/C<sub>70</sub><sub>(1-x)</sub> fullerene mixture*. In: *Phys. Solid State* 57.2 (2015), pp. 412–418.
- [133] C. Rey, L. J. Gallego, and J. A. Alonso. *Molecular-dynamics study of the structure, binding energy, and melting of small clusters of fullerene molecules using Girifalco's spherical model*. In: *Phys. Rev. B* 49.12 (Mar. 1994), pp. 8491–8494.
- [134] W. Jin et al. *C<sub>60</sub> chain phases on ZnPc/Ag(111) surfaces: Supramolecular organization driven by competing interactions*. In: *J. Chem. Phys.* 142.10 (Mar. 2015), p. 101910.

- [135] C. Patrick Royall and Stephen R. Williams. *C<sub>60</sub>: The First One-Component Gel?* In: *J. Phys. Chem. B* 115.22 (June 2011), pp. 7288–7293.
- [136] Liang Hu et al. *Mixed Layered Growth of Fullerene C<sub>60</sub> Self-Assembly on an Oxygen-Passivated Fe(001)-p(1 × 1)O Surface.* In: *J. Phys. Chem. C* 123.25 (June 2019), pp. 15477–15482.
- [137] M. C. Abramo et al. *Atomistic versus two-body central potential models of C<sub>60</sub>: A comparative molecular dynamics study.* In: *Phys. Rev. E* 69.3 (Mar. 2004), p. 031112.
- [138] P. A. Gravil et al. *Adsorption of C<sub>60</sub> molecules.* In: *Phys. Rev. B* 53 (3 1996), pp. 1622–1629.
- [139] Joseph E. Goose, Eric L. First, and Paulette Clancy. *Nature of step-edge barriers for small organic molecules.* In: *Phys. Rev. B* 81 (20 2010), p. 205310.
- [140] M. Schunack et al. *Long Jumps in the Surface Diffusion of Large Molecules.* In: *Phys. Rev. Lett.* 88.15 (Mar. 2002), p. 156102.
- [141] Jens Weckesser, Johannes V. Barth, and Klaus Kern. *Direct observation of surface diffusion of large organic molecules at metal surfaces: PVBA on Pd(110).* In: *J. Chem. Phys.* 110.11 (Mar. 1999), pp. 5351–5354.
- [142] Christian Maes and Karel Netočný. In: *Journal of Statistical Physics* 110.1/2 (2003), pp. 269–310.
- [143] Udo Seifert. *Entropy Production along a Stochastic Trajectory and an Integral Fluctuation Theorem.* In: *Phys. Rev. Lett.* 95 (4 2005), p. 040602.
- [144] Sebastian Bommel. “Unravelling nanoscale molecular processes in organic thin films”. PhD thesis. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 2015.
- [145] Y. Kuwahara. *Muscovite surface structure imaged by fluid contact mode AFM.* In: *Phys. Chem. Miner.* 26.3 (Feb. 1999), pp. 198–205.
- [146] Takeshi Fukuma et al. *True atomic resolution in liquid by frequency-modulation atomic force microscopy.* In: *Appl. Phys. Lett.* 87.3 (July 2005), p. 034101.
- [147] Hendrik Heinz et al. *Force Field for Mica-Type Silicates and Dynamics of Octadecylammonium Chains Grafted to Montmorillonite.* In: *Chem. Mater.* 17.23 (Nov. 2005), pp. 5658–5669.
- [148] F. Ostendorf et al. *How flat is an air-cleaved mica surface?* In: *Nanotechnology* 19.30 (June 2008), p. 305705.
- [149] Saulius Gražulis et al. *Crystallography Open Database – an open-access collection of crystal structures.* In: *J. Appl. Crystallogr.* 42.4 (May 2009), pp. 726–729.



- [150] M. J. Gillan. *Collective dynamics in superionic CaF<sub>2</sub>. I. Simulation compared with neutron-scattering experiment*. In: *J. Phys. C* 19.18 (June 1986), pp. 3391–3411.
- [151] Nora H. de Leeuw and Timothy G. Cooper. *A computational study of the surface structure and reactivity of calcium fluoride*. In: *J. Mater. Chem.* 13 (1 2003), pp. 93–101.
- [152] J. Breton, J. Gonzalez-Platas, and C. Girardet. *Endohedral and exohedral adsorption in C<sub>60</sub>: An analytical model*. In: *J. Chem. Phys.* 99.5 (Sept. 1993), pp. 4036–4040.
- [153] M. C. Abramo and C. Caccamo. *A molecular dynamics study of impurity desorption from solid clusters of rigid C<sub>60</sub> molecules*. In: *J. Chem. Phys.* 106.15 (Apr. 1997), pp. 6475–6482.
- [154] S. Palucha, K. Kaczor, and Z. Gburski. *A molecular dynamics study of a fullerene-cyanoadamantane mixture*. In: *J. Phys. Condens. Matter* 14.6 (Jan. 2002), pp. 1223–1230.
- [155] Wolfram Research, Inc. *Mathematica, Version 12.1*. 2020.
- [156] J. Klafter and G. Zumofen. *Lévy statistics in a Hamiltonian system*. In: *Phys. Rev. E* 49.6 (June 1994), pp. 4873–4877.
- [157] Joseph Klafter, Michael F. Shlesinger, and Gert Zumofen. *Beyond Brownian Motion*. In: *Phys. Today* 49.2 (Feb. 1996), pp. 33–39.
- [158] Charles F.F. Karney. *Quaternions in molecular modeling*. In: *J. Mol. Graph. Model.* 25.5 (Jan. 2007), pp. 595–604.
- [159] Pu Liu, Dimitris K. Agrafiotis, and Douglas L. Theobald. *Fast determination of the optimal rotational matrix for macromolecular superpositions*. In: *J. Comput. Chem.* 31.7 (2009), pp. 1561–1563.
- [160] Mehdi Jafary-Zadeh et al. *Kinetic nanofriction: a mechanism transition from quasi-continuous to ballistic-like Brownian regime*. In: *Nanoscale Res. Lett.* 7.1 (2012), p. 148.
- [161] Martin Körner. “Analyse und Entwicklung kinetischer Modelle für das Clusterwachstum auf Oberflächen”. PhD thesis. 2012.
- [162] Yael S. Elmatad, David Chandler, and Juan P. Garrahan. *Corresponding States of Structural Glass Formers*. In: *J. Phys. Chem. B* 113.16 (Mar. 2009), pp. 5563–5567.
- [163] Yael S. Elmatad, David Chandler, and Juan P. Garrahan. *Corresponding States of Structural Glass Formers. II*. In: *J. Phys. Chem. B* 114.51 (Dec. 2010), pp. 17113–17119.
- [164] Mads Haahr. URL: <https://www.random.org> (visited on 05/29/2021).

- [165] Makoto Matsumoto and Takuji Nishimura. *Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator*. In: *ACM T. Model. Comput. S.* 8.1 (Jan. 1998), pp. 3–30.
- [166] Mark Galassi. *GNU Scientific Library - Reference Manual*. Network Theory, 2009.
- [167] Achim Klenke. *Wahrscheinlichkeitstheorie*. Springer Berlin Heidelberg, 2020.

## B.5 ACKNOWLEDGEMENTS

(deleted from the electronic version of this dissertation)



## CURRICULUM VITAE

(deleted from the electronic version of this dissertation)